GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 19:50:48 ; Search time 159 Seconds
(without alignments)
2058.090 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817199/runat_18062003_145103_857/app_query.fasta_1.391
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09817199_@CGN_1_1_57_@runat_18062003_145103_857
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1150	100.0	1106	9	US-10-037-270-959	Sequence 959, App
	2	1150	100.0	2674	10	US-09-817-199A-1	Sequence 1, Appli
	3	1145	99.6	1116	10	US-09-794-257-13	Sequence 13, Appl
	4	1144	99.5	2623	9	US-09-764-868-71	Sequence 71, Appl
	5	1140	99.1	875	12	US-10-051-986-10	Sequence 10, Appl
	6	1133	98.5	1316	9	US-09-764-868-493	Sequence 493, App
	7	977	85.0	576	10	US-09-794-257-15	Sequence 15, Appl
	8	720	62.6	447	10	US-09-867-550-1811	Sequence 1811, Ap
	9	540	47.0	434	9	US-09-918-995-648	Sequence 648, App
	10	535	46.5	964	9	US-09-764-868-75	Sequence 75, Appl
	11	511.5	44.5	2497	10	US-09-834-975-879	Sequence 879, App
	12	511.5	44.5	2497	10	US-09-834-975-885	Sequence 885, App
	13	511.5	44.5	2497	10	US-09-834-975-894	Sequence 894, App
	14	511.5	44.5	2497	10	US-09-834-975-896	Sequence 896, App
	15	507	44.1	881	10	US-09-770-445-529	Sequence 529, App
	16	506	44.0	1537	10	US-09-925-300-631	Sequence 631, App
	17	501	43.6	624	10	US-09-794-257-9	Sequence 9, Appli
	18	501	43.6	1161	10	US-09-794-257-7	Sequence 7, Appli
	19	500	43.5	651	9	US-09-938-842A-836	Sequence 836, App
	20	496.5	43.2	609	9	US-09-938-842A-832	Sequence 832, App
	21	483	42.0	1274	10	US-09-925-302-91	Sequence 91, Appl
	22	470	40.9	925	10	US-09-967-736-4	Sequence 4, Appli
	23	469.5	40.8	939	9	US-10-102-806-48	Sequence 48, Appl
	24	448.5	39.0	639	10	US-09-350-874-66	Sequence 66, Appl
	25	446	38.8	833	9	US-10-037-270-426	Sequence 426, App
	26	435	37.8	2021	9	US-09-764-868-88	Sequence 88, Appl
	27	435	37.8	3257	10	US-09-817-198A-1	Sequence 1, Appli
	28	432	37.6	1000	9	US-09-764-868-488	Sequence 488, App
	29	429.5	37.3	566	9	US-09-764-868-507	Sequence 507, App
	30	420.5	36.6	896	10	US-09-770-445-478	Sequence 478, App
c	31	413.5	36.0	1127	9	US-09-764-868-487	Sequence 487, App
	32	413.5	36.0	3124	10	US-09-925-302-340	Sequence 340, App
	33	413	35.9	2873	9	US-10-198-846-11497	Sequence 11497, A
c	34	411	35.7	894	10	US-09-770-445-487	Sequence 487, App
	35	408	35.5	1061	10	US-09-880-107-3393	Sequence 3393, Ap
	36	403.5	35.1	1069	9	US-10-037-270-646	Sequence 646, App
	37	401.5	34.9	585	10	US-09-917-800A-1461	Sequence 1461, Ap
	38	399.5	34.7	771	10	US-09-728-445-652	Sequence 652, App
	39	392.5	34.1	654	9	US-09-938-842A-2113	Sequence 2113, Ap
	40	388	33.7	725	9	US-10-198-846-13439	Sequence 13439, A
	41	386	33.6	2771	10	US-09-834-765-1	Sequence 1, Appli
	42	385	33.5	857	10	US-09-917-800A-1426	Sequence 1426, Ap
	43	383	33.3	514	9	US-09-911-904-150	Sequence 150, App
	44	383	33.3	538	10	US-09-924-035A-794	Sequence 794, App
	45	383	33.3	1042	10	US-09-954-456-563	Sequence 563, App

ALIGNMENTS

RESULT 1
US-10-037-270-959
; Sequence 959, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317

```
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 959
;   LENGTH: 1106
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)..(699)
US-10-037-270-959

Alignment Scores:
Pred. No.:              4.31e-139      Length:          1106
Score:                  1150.00        Matches:         223
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     9              Gaps:            0

US-09-817-199B-2 (1-223) x US-10-037-270-959 (1-1106)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 87

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 207

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 267

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 327

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 387

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 447

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 507

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 567

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 627

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 687

Qy        221 SerPheMet 223
              |||||||||
Db        688 TCCTTCATG 696

RESULT 2
```

```
US-09-817-199A-1
; Sequence 1, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001187
;  CURRENT APPLICATION NUMBER: US/09/817,199A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2674
;   TYPE: DNA
;   ORGANISM: Human
US-09-817-199A-1

Alignment Scores:
Pred. No.:              1.6e-138       Length:          2674
Score:                  1150.00        Matches:         223
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     10             Gaps:            0

US-09-817-199B-2 (1-223) x US-09-817-199A-1 (1-2674)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 101

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 161

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 221

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 281

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 341

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 401

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 461

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 521

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 581

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 641

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
```

```
Qy        221 SerPheMet 223
              |||||||||
Db        702 TCCTTCATG 710

RESULT 3
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 1116
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)...(699)
US-09-794-257-13

Alignment Scores:
Pred. No.:              1.94e-138      Length:          1116
Score:                  1145.00        Matches:         222
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            99.57%         Indels:          0
DB:                     10             Gaps:            0

US-09-817-199B-2 (1-223) x US-09-794-257-13 (1-1116)

Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 90

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 150

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 210

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 270

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 330

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 390

Qy        122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 450

Qy        142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 510

Qy        162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 570

Qy        182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 630

Qy        202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 690

Qy        222 PheMet 223
              ||||||
Db        691 TTCATG 696

RESULT 4
US-09-764-868-71
; Sequence 71, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
;   LENGTH: 2623
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-868-71

Alignment Scores:
Pred. No.:              9.31e-138      Length:          2623
Score:                  1144.00        Matches:         222
Percent Similarity:     99.55%         Conservative:    0
Best Local Similarity:  99.55%         Mismatches:      1
Query Match:            99.48%         Indels:          0
DB:                     9              Gaps:            0

US-09-817-199B-2 (1-223) x US-09-764-868-71 (1-2623)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 71

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 131

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 191

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        192 GCCACCGTCGGCATAGACTTCAGCAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 251

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 311

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 371

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 431

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 491
```

```
Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 551

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 611

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 671

Qy        221 SerPheMet 223
              |||||||||
Db        672 TCCTTCATG 680

RESULT 5
US-10-051-986-10
; Sequence 10, Application US/10051986
; Patent No. US20020146770A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Tang, Y. Tom
;                   Lal, Preeti
;                   Guegler, Karl J.
;                   Corley, Neil C.
;                   Patterson, Chandra
;                   Batra, Sajeev
;                   Baughn, Mariah R.
;        TITLE OF INVENTION: RAS PROTEINS
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/051,986
;             FILING DATE: 15-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/766,551
;             FILING DATE: DECEMBER 12, 1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C.
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 875 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: UCMCL5T01
;             CLONE: 1528559
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10 :
US-10-051-986-10

Alignment Scores:
Pred. No.:              6.01e-138       Length:         875
Score:                  1140.00         Matches:        221
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            99.13%          Indels:         0
DB:                     12              Gaps:           0

US-09-817-199B-2 (1-223) x US-10-051-986-10 (1-875)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProProCys 22
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGC 62

Qy         23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAA 122

Qy         43 ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACC 182

Qy         63 ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAG 242

Qy         83 IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 ATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGAT 302

Qy        103 AlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGG 362

Qy        123 AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGC 422

Qy        143 AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCC 482

Qy        163 ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTA 542

Qy        183 AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCC 602

Qy        203 SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTC 662

Qy        223 Met 223
              |||
Db        663 ATG 665

RESULT 6
US-09-764-868-493
; Sequence 493, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 493
;   LENGTH: 1316
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1281)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1299)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-493

Alignment Scores:
Pred. No.:              8.88e-137      Length:          1316
Score:                  1133.00        Matches:         220
Percent Similarity:     99.10%         Conservative:    0
Best Local Similarity:  99.10%         Mismatches:      2
Query Match:            98.52%         Indels:          0
DB:                     9              Gaps:            0

US-09-817-199B-2 (1-223) x US-09-764-868-493 (1-1316)

Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 60

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 120

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 180

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGTCGGCATAGACTTCAGSAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 240

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 300

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 360

Qy        122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 420

Qy        142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 480

Qy        162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAGGGAGTACGGKGTTCCCTTYCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 540

Qy        182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 600

Qy        202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        601 CCCAGCTTCCAGATCCGAGACTATGTAKAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 660

Qy        222 PheMet 223
              ||||||
Db        661 TTCATG 666

RESULT 7
US-09-794-257-15
; Sequence 15, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 576
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-794-257-15

Alignment Scores:
Pred. No.:              4.15e-117      Length:          576
Score:                  977.00         Matches:         191
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            84.96%         Indels:          0
DB:                     10             Gaps:            0

US-09-817-199B-2 (1-223) x US-09-794-257-15 (1-576)

Qy         33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60

Qy         53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy         73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy         93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIle 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        213 GlnLysLysArgSerSerCysCysSerPheMet 223
              ||||||||||||||||||||||||||||||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATG 573

RESULT 8
US-09-867-550-1811
; Sequence 1811, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
```

```
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1811
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)
;   OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1811

Alignment Scores:
Pred. No.:                5.47e-84        Length:         447
Score:                    720.00          Matches:        138
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              62.61%          Indels:         0
DB:                       10              Gaps:           0

US-09-817-199B-2 (1-223) x US-09-867-550-1811 (1-447)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 93

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 153

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 213

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 273

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 333

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 393

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIle 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATC 447

RESULT 9
US-09-918-995-648
; Sequence 648, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 648
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(434)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-648

Alignment Scores:
Pred. No.:                1.07e-60        Length:         434
Score:                    540.00          Matches:        99
Percent Similarity:       93.39%          Conservative:   14
Best Local Similarity:    81.82%          Mismatches:     8
Query Match:              46.96%          Indels:         0
DB:                       9               Gaps:           0

US-09-817-199B-2 (1-223) x US-09-918-995-648 (1-434)

Qy         26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
              ||||||::      ||||||||||||:::|||||||:::||||||||||||||||||||
Db         67 TACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTG 126

Qy         46 LeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIle 65
              |||:::::::||||||||||||||||||||||:::||||||||||||:::||||||||||||
Db        127 CTGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATT 186

Qy         66 AspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAsp 85
              ||||||||||||||||||:::   ||||||||||||:::||||||||||||:::|||||||
Db        187 GACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGAC 246

Qy         86 ThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAla 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        247 ACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCT 306

Qy        106 LeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeu 125
              ||||||||||||||||||:::|||||||||:::||||||||||||||:::|||||||||
Db        307 CTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTG 366

Qy        126 ThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAla 145
              ||||||||||||||||||||||||   ||||||   :::||||||||||||||||||||
Db        367 ACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTG 426

Qy        146 Asp 146
              |||
Db        427 GAC 429

RESULT 10
US-09-764-868-75
; Sequence 75, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
;   LENGTH: 964
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (806)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (898)
;   OTHER INFORMATION: n equals a,t,g, or c
```

```
;   NAME/KEY: SITE
;   LOCATION: (918)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (924)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (952)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (959)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-75

Alignment Scores:
Pred. No.:              1.56e-59        Length:         964
Score:                  535.00          Matches:        98
Percent Similarity:     85.40%          Conservative:   19
Best Local Similarity:  71.53%          Mismatches:     20
Query Match:            46.52%          Indels:         0
DB:                     9               Gaps:           0

US-09-817-199B-2 (1-223) x US-09-764-868-75 (1-964)

Qy         84 TrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAla 103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TGGGACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCT 62

Qy        104 GlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAla 123
              ||||||||||||||||||||:::|||||||||:::||||||||||||||||||:::|||
Db         63 CATGCTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCC 122

Qy        124 TrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsn 143
              |||||||||||||||||||||||||||||   ||||||   :::|||||||||||||||
Db        123 TGGCTGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAAC 182

Qy        144 LysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArg 163
              |||   |||   :::   |||||||||::::::   ||||||||||||   ||||||:::
Db        183 AAGGTGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAG 242

Qy        164 GluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAla 183
              |||||||||:::||||||:::||||||||||||||||||||||:::||||||:::||||||
Db        243 GAGTATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCC 302

Qy        184 PheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSer 203
              |||   ||||||||||||||||||||   |||:::                 ||||||
Db        303 TTCACAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGC 362

Qy        204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              |||::::::   |||||||||:::   :::   :::   :::|||||||||
Db        363 TTCCGGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 413

RESULT 11
US-09-834-975-879
; Sequence 879, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Brown, Jeffrey
;  APPLICANT:  Bolt, Andrew
;  APPLICANT:  Van Huffel, Christophe
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;  TITLE OF INVENTION:  FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF HUMAN CANCERS
;  FILE REFERENCE: MRI-016B
;  CURRENT APPLICATION NUMBER: US/09/834,975
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 60/197,538
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 1046
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 879
;   LENGTH: 2497
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2497)
;   OTHER INFORMATION: n = A,T,C or G
US-09-834-975-879

Alignment Scores:
Pred. No.:              7.09e-56        Length:         2497
Score:                  511.50          Matches:        106
Percent Similarity:     62.82%          Conservative:   41
Best Local Similarity:  45.30%          Mismatches:     68
Query Match:            44.48%          Indels:         19
DB:                     10              Gaps:           4

US-09-817-199B-2 (1-223) x US-09-834-975-879 (1-2497)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
              |||||||||              |||             |||      |||||||||
Db          8 GGCACGCCC------------CCTCGCCCGCGGCCCCCTCCCCGCCTCTCTCCACCGCCT 55

Qy         22 ----------------------------------CysSerProSerTyrAspLeuThr 29
                                                   :::   :::||||||
Db         56 CCTCTGGCTCCCCGGTCAGAGGGCCGGAGCGAGAAGATGGCGAAGACGTACGATTATCTC 115

Qy         30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
                |||:::::|||:::||||||:::||||||||||||||||||   |||   :::|||
Db        116 TTCAAGCTCCTGCTGATCGGCGACTCGGGGGTAGGCAAGACCTGCCTCCTGTTCCGCTTC 175

Qy         50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
                :::   ||||||   :::   ||||||||||:::|||:::||||||||||:::
Db        176 TCAGAGGACGCCTTC---AACACCACCTTCATCTCCACCATCGGAATTGATTTTAAAATT 232

Qy         70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
              :::   :::   :::|||||   ::::::||||||||||||||||||||||||||||||
Db        233 AGAACGATAGAACTAGATGGAAAGAAAATTAAGCTTCAGATATGGGACACAGCGGGTCAG 292

Qy         90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeu 109
              ||||||||||||::::::|||   ||||||||||   |||      ::::::|||:::
Db        293 GAAAGATTCCGAACAATCACGACAGCGTACTACAGAGGAGCCATGGGCATTATGCTGGTC 352

Qy        110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
              |||||||||||||:::   |||||||||||||:::   |||:::      |||
Db        353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAAAAATTGGATCAGAAACATTGAA 412

Qy        130 GluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
              |||:::|||      ||||||      |||:::||||||||||||   ||||||:::
Db        413 GAGCATGCCTCTTCCGATGTCGAAAGAATGATCCTGGGTAACAAATGTGATATGAATGAC 472

Qy        150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
              :::|||   :::      |||   ||||||   ||||||   :::|||||:::   |||
Db        473 AAAAGACAAGTGTCAAAAGAAAGAGGGGAGAAGCTAGCAATTGACTATGGGATTAAATTC 532

Qy        170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
              ||||||||||||||||:::      |||||||||   ||||||      :::|||:::
Db        533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTTTACACTTGCACGA 592

Qy        190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
              ::::::      :::         :::      |||      |||
Db        593 GATATAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCAGCAGGAGCAGGTGGA 652

Qy        209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
              |||:::      :::   ||||||         ||||||
Db        653 CCAGTGAAAATAACAGAAAACCGATCAAAGAAGACCAGTTTC 694

RESULT 12
US-09-834-975-885
; Sequence 885, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Brown, Jeffrey
```

```
;  APPLICANT:  Bolt, Andrew
;  APPLICANT:  Van Huffel, Christophe
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;  TITLE OF INVENTION:  FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF HUMAN CANCERS
;  FILE REFERENCE: MRI-016B
;  CURRENT APPLICATION NUMBER: US/09/834,975
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 60/197,538
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 1046
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 885
;   LENGTH: 2497
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2497)
;   OTHER INFORMATION: n = A,T,C or G
US-09-834-975-885

Alignment Scores:
Pred. No.:              7.09e-56      Length:          2497
Score:                  511.50        Matches:         106
Percent Similarity:     62.82%        Conservative:    41
Best Local Similarity:  45.30%        Mismatches:      68
Query Match:            44.48%        Indels:          19
DB:                     10            Gaps:            4

US-09-817-199B-2 (1-223) x US-09-834-975-885 (1-2497)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
              |||||||||              |||             |||      |||||||||
Db          8 GGCACGCCC------------CCTCGCCCGCGGCCCCCTCCCCGCCTCTCTCCACCGCCT 55

Qy         22 -----------------------------------CysSerProSerTyrAspLeuThr 29
                                                    :::   :::||||||
Db         56 CCTCTGGCTCCCCGGTCAGAGGGCCGGAGCGAGAAGATGGCGAAGACGTACGATTATCTC 115

Qy         30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
                |||:::::|||:::||||||:::||||||||||||||||||||   |||   :::|||
Db        116 TTCAAGCTCCTGCTGATCGGCGACTCGGGGGTAGGCAAGACCTGCCTCCTGTTCCGCTTC 175

Qy         50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
                 :::   ||||||   :::   |||||||||:::|||:::|||||||||||||:::
Db        176 TCAGAGGACGCCTTC---AACACCACCTTCATCTCCACCATCGGAATTGATTTTAAAATT 232

Qy         70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
                 :::   :::   :::||||||   ::::::||||||||||||||||||||||||||||
Db        233 AGAACGATAGAACTAGATGGAAAGAAAATTAAGCTTCAGATATGGGACACAGCGGGTCAG 292

Qy         90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeu 109
              ||||||||||||:::::|||   ||||||||||||   |||      :::::::|||:::
Db        293 GAAAGATTCCGAACAATCACGACAGCGTACTACAGAGGAGCCATGGGCATTATGCTGGTC 352

Qy        110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
              ||||||||||||||:::   ||||||||||||||:::   |||:::      |||
Db        353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAAAAATTGGATCAGAAACATTGAA 412

Qy        130 GluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
              |||:::|||      ||||||      |||:::||||||||||||   ||||||:::
Db        413 GAGCATGCCTCTTCCGATGTCGAAAGAATGATCCTGGGTAACAAATGTGATATGAATGAC 472

Qy        150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
              :::|||   :::      |||   ||||||   ||||||   :::|||||:::   |||
Db        473 AAAAGACAAGTGTCAAAAGAAAGAGGGGAGAAGCTAGCAATTGACTATGGGATTAAATTC 532

Qy        170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
              ||||||||||||||||:::      |||||||||   ||||||      :::|||:::
Db        533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTTTACACTTGCACGA 592

Qy        190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
                                                  :::::      :::      :::      |||      |||
Db        593 GATATAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCAGCAGGAGCAGGTGGA 652

Qy        209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
                |||:::      :::   ||||||         |||||
Db        653 CCAGTGAAAATAACAGAAAACCGATCAAAGAAGACCAGTTTC 694

RESULT 13
US-09-834-975-894
; Sequence 894, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Brown, Jeffrey
;  APPLICANT:  Bolt, Andrew
;  APPLICANT:  Van Huffel, Christophe
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;  TITLE OF INVENTION:  FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF HUMAN CANCERS
;  FILE REFERENCE: MRI-016B
;  CURRENT APPLICATION NUMBER: US/09/834,975
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 60/197,538
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 1046
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 894
;   LENGTH: 2497
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2497)
;   OTHER INFORMATION: n = A,T,C or G
US-09-834-975-894

Alignment Scores:
Pred. No.:              7.09e-56      Length:          2497
Score:                  511.50        Matches:         106
Percent Similarity:     62.82%        Conservative:    41
Best Local Similarity:  45.30%        Mismatches:      68
Query Match:            44.48%        Indels:          19
DB:                     10            Gaps:            4

US-09-817-199B-2 (1-223) x US-09-834-975-894 (1-2497)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
              |||||||||              |||             |||      |||||||||
Db          8 GGCACGCCC------------CCTCGCCCGCGGCCCCCTCCCCGCCTCTCTCCACCGCCT 55

Qy         22 -----------------------------------CysSerProSerTyrAspLeuThr 29
                                                    :::   :::||||||
Db         56 CCTCTGGCTCCCCGGTCAGAGGGCCGGAGCGAGAAGATGGCGAAGACGTACGATTATCTC 115

Qy         30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
                |||:::|::|||:::||||||:::||||||||||||||||||   |||   :::|||
Db        116 TTCAAGCTCCTGCTGATCGGCGACTCGGGGGTAGGCAAGACCTGCCTCCTGTTCCGCTTC 175

Qy         50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
                 :::   ||||||   :::   |||||||||:::|||:::|||||||||||||:::
Db        176 TCAGAGGACGCCTTC---AACACCACCTTCATCTCCACCATCGGAATTGATTTTAAAATT 232

Qy         70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
                 :::   :::   :::||||||   ::::::||||||||||||||||||||||||||||
Db        233 AGAACGATAGAACTAGATGGAAAGAAAATTAAGCTTCAGATATGGGACACAGCGGGTCAG 292

Qy         90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeu 109
              ||||||||||||:::::|||   ||||||||||||   |||      :::::|||:::
Db        293 GAAAGATTCCGAACAATCACGACAGCGTACTACAGAGGAGCCATGGGCATTATGCTGGTC 352

Qy        110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
              ||||||||||||||:::   ||||||||||||||:::   |||:::      |||
Db        353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAAAAATTGGATCAGAAACATTGAA 412
```

```
Qy        130 GluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
              |||:::|||      ||||||      |||:::||||||||||||   ||||||:::
Db        413 GAGCATGCCTCTTCCGATGTCGAAAGAATGATCCTGGGTAACAAATGTGATATGAATGAC 472

Qy        150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
              :::|||   :::      |||   ||||||   ||||||   :::|||||:::   |||
Db        473 AAAAGACAAGTGTCAAAAGAAAGAGGGGAGAAGCTAGCAATTGACTATGGGATTAAATTC 532

Qy        170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
              ||||||||||||||||||:::      |||||||||   ||||||      :::|||:::
Db        533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTTTACACTTGCACGA 592

Qy        190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
              ::::::      :::          :::      |||      |||
Db        593 GATATAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCAGCAGGAGCAGGTGGA 652

Qy        209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
                 |||:::      :::   ||||||         ||||||
Db        653 CCAGTGAAAATAACAGAAAACCGATCAAAGAAGACCAGTTTC 694

RESULT 14
US-09-834-975-896
; Sequence 896, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Brown, Jeffrey
;  APPLICANT:  Bolt, Andrew
;  APPLICANT:  Van Huffel, Christophe
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;  TITLE OF INVENTION:  FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF HUMAN CANCERS
;  FILE REFERENCE: MRI-016B
;  CURRENT APPLICATION NUMBER: US/09/834,975
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 60/197,538
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 1046
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 896
;   LENGTH: 2497
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2497)
;   OTHER INFORMATION: n = A,T,C or G
US-09-834-975-896

Alignment Scores:
Pred. No.:              7.09e-56        Length:         2497
Score:                  511.50          Matches:        106
Percent Similarity:     62.82%          Conservative:   41
Best Local Similarity:  45.30%          Mismatches:     68
Query Match:            44.48%          Indels:         19
DB:                     10              Gaps:           4

US-09-817-199B-2 (1-223) x US-09-834-975-896 (1-2497)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
              ||||||||                |||          |||      |||||||||
Db          8 GGCACGCCC------------CCTCGCCCGCGGCCCCCTCCCCGCCTCTCTCCACCGCCT 55

Qy         22 ----------------------------------CysSerProSerTyrAspLeuThr 29
                                               :::   :::||||||
Db         56 CCTCTGGCTCCCCGGTCAGAGGGCCGGAGCGAGAAGATGGCGAAGACGTACGATTATCTC 115

Qy         30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
                 |||:::::::|||:::|||||||:::|||||||||||||||||||   |||   :::|||
Db        116 TTCAAGCTCCTGCTGATCGGCGACTCGGGGGTAGGCAAGACCTGCCTCCTGTTCCGCTTC 175

Qy         50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
                 :::   ||||||   :::   |||||||||:::|||:::||||||||||||:::
Db        176 TCAGAGGACGCCTTC---AACACCACCTTCATCTCCACCATCGGAATTGATTTTAAAATT 232

Qy         70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
              :::   :::   :::||||||   ::::::|||||||||||||||||||||||||||||
Db        233 AGAACGATAGAACTAGATGGAAAGAAAATTAAGCTTCAGATATGGGACACAGCGGGTCAG 292

Qy         90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeu 109
              ||||||||||||:::::::|||   ||||||||||||   |||      ::::::|||:::
Db        293 GAAAGATTCCGAACAATCACGACAGCGTACTACAGAGGAGCCATGGGCATTATGCTGGTC 352

Qy        110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
              ||||||||||||||||:::   ||||||||||||||||:::   |||:::      |||
Db        353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAAAAATTGGATCAGAAACATTGAA 412

Qy        130 GluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
              |||:::|||      ||||||      |||:::||||||||||||   ||||||:::
Db        413 GAGCATGCCTCTTCCGATGTCGAAAGAATGATCCTGGGTAACAAATGTGATATGAATGAC 472

Qy        150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
              :::|||   :::      |||   ||||||   ||||||   :::|||||:::   |||
Db        473 AAAAGACAAGTGTCAAAAGAAAGAGGGGAGAAGCTAGCAATTGACTATGGGATTAAATTC 532

Qy        170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
              ||||||||||||||||||:::      |||||||||   ||||||      :::|||:::
Db        533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTTTACACTTGCACGA 592

Qy        190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
              ::::::      :::          :::      |||      |||
Db        593 GATATAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCAGCAGGAGCAGGTGGA 652

Qy        209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
                 |||:::      :::   ||||||         ||||||
Db        653 CCAGTGAAAATAACAGAAAACCGATCAAAGAAGACCAGTTTC 694

RESULT 15
US-09-770-445-529
; Sequence 529, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2023US (PARA-012PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,445
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  US 60/178,472
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 529
;   LENGTH: 881
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-770-445-529
```

Alignment Scores:

Pred. No.:	5.77e-56	Length:	881
Score:	507.00	Matches:	95
Percent Similarity:	69.38%	Conservative:	50
Best Local Similarity:	45.45%	Mismatches:	56
Query Match:	44.09%	Indels:	8
DB:	10	Gaps:	3

US-09-817-199B-2 (1-223) x US-09-770-445-529 (1-881)

```
Qy         18 ArgSerProProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAsp 37
              ||||||         :::|||   ||||||         |||:::::::|||:::||||||
Db         48 AGATCGAAAACCATGAATCCTGAGTACGACTATCTTTTCAAGCTCCTGCTTATCGGGGAT 107

Qy         38 ThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGly 57
              :::|||||||||||||:::|||   |||:::::::|||   |||   :::::::::
Db        108 TCTGGCGTAGGCAAGTCTTGTCTTCTTTTGAGATTCTCTGATGATTCTTATGTAGAA--- 164

Qy         58 ThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyVal 77
              ::::::|||:::|||:::|||:::||||||:::   :::   |||      ||||||
Db        165 AGTTACATTAGCACTATTGGAGTCGATTTTAAAATTAGGACTGTGGAACAAGATGGCAAA 224

Qy         78 ArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHis 97
              :::||||||||||||||||||||||||||||||||||||||||||:::::|||
Db        225 ACAATTAAGCTCCAAATTTGGGACACTGCTGGTCAAGAACGGTTCAGGACTATTACTAGC 284

Qy         98 AlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSer 117
              :::|||||||||   |||      ::::::::::::||||||:::|||:::::   |||
Db        285 AGTTACTACCGTGGGGCACATGGAATTATTATTGTCTACGATGTCACAGATGAAGAAAGC 344

Qy        118 PheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValVal 137
              |||:::|||:::::   ||||||:::||||||      ||||||      :::|||
Db        345 TTCAATAATGTCAAGCAATGGTTGAGTGAAATTGATCGTTATGCTAGTGACAATGTCAAC 404

Qy        138 IleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAsp 157
              :::|||:::|||||||||:::|||::::::      |||   |||      |||
Db        405 AAACTCCTTGTTGGAAACAAGTCTGATCTTACTGAAAACAGAGCCATTCCTTATGAAACT 464

Qy        158 GlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGly 177
              :::      |||   |||   |||:::||||||:::||||||||||||||||
Db        465 GCCAAGGCTTTTGCCGATGAAATCGGGATTCCTTTTATGGAGACTAGTGCAAAAGATGCT 524

Qy        178 MetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg--------- 194
              |||||||||   ||||||:::|||::::::      :::|||   |||
Db        525 ACAAACGTAGAACAGGCTTTCATGGCAATGTCTGCATCCATCAAAGAGAGAATGGCTAGC 584

Qy        195 ------AlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
                   ||||||:::   |||      |||:::   |||||||||           :::
Db        585 CAACCAGCTGGGAATAATGCAAGACCACCGACCGTGCAGATCAGAGGACAGCCTGTGGCA 644

Qy        213 GlnLysLysArgSerSerCysCysSer 221
              ||||||      :::   |||||||||
Db        645 CAGAAG------AACGGCTGCTGCTCA 665
```

Search completed: June 23, 2003, 21:35:26
Job time : 163 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 19:39:18 ; Search time 1382 Seconds
(without alignments)
2613.309 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817199/runat_18062003_145102_801/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199_@CGN_1_1_1906_@runat_18062003_145102_801 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	1014	14	BM921365	BM921365 AGENCOURT
2	1137	98.9	1051	12	BG283602	BG283602 602407930
3	1134	98.6	887	9	AL522282	AL522282 AL522282
4	1058	92.0	676	10	BB598938	BB598938 BB598938
5	1020	88.7	651	10	BB633978	BB633978 BB633978
6	1006	87.5	740	13	BI767046	BI767046 603054267
7	1001	87.0	655	10	BB635649	BB635649 BB635649
8	856.5	74.5	749	9	AL559085	AL559085 AL559085
9	856	74.4	733	10	BB619625	BB619625 BB619625
10	852	74.1	522	13	BM151643	BM151643 TCBAP1E10
c 11	829	72.1	736	13	BG934396	BG934396 SK1-0698
12	825	71.7	676	10	BB206788	BB206788 BB206788
13	790	68.7	553	12	BE755280	BE755280 209068 MA
14	763	66.3	488	13	BM149118	BM149118 TCAAP2E63
15	760	66.1	888	14	BQ687186	BQ687186 AGENCOURT
16	760	66.1	953	12	BE906081	BE906081 601497118
17	754	65.6	895	14	BQ688147	BQ688147 AGENCOURT
18	752	65.4	579	10	BB621927	BB621927 BB621927
19	740	64.3	657	10	BB196489	BB196489 BB196489
20	739	64.3	455	10	AW464449	AW464449 BP230015B
21	728	63.3	691	12	BG253976	BG253976 602366910
22	721.5	62.7	672	13	BM492040	BM492040 pgp2n.pk0
23	650	56.5	701	10	AW914157	AW914157 EST345461
24	646	56.2	708	9	AJ442405	AJ442405 AJ442405
25	622	54.1	780	12	BG784310	BG784310 SEAUMC004
26	618	53.7	443	10	AW956878	AW956878 EST368948
27	610	53.0	380	10	BE241772	BE241772 TCAAP2E02
28	607	52.8	386	10	AW484587	AW484587 61493 MAR
29	588	51.1	820	13	BI916873	BI916873 603177832
30	582	50.6	782	13	BI545231	BI545231 603187353
31	581.5	50.6	658	10	AV697217	AV697217 AV697217
32	581	50.5	368	13	BM152052	BM152052 TCBAP1E12
33	580.5	50.5	648	10	AV697216	AV697216 AV697216
34	580.5	50.5	684	10	AV692701	AV692701 AV692701
35	580.5	50.5	697	10	AV692689	AV692689 AV692689
36	576.5	50.1	678	10	AV685389	AV685389 AV685389
37	572	49.7	668	10	AV688197	AV688197 AV688197
38	557.5	48.5	647	10	AV698293	AV698293 AV698293
39	557	48.4	678	10	AV686507	AV686507 AV686507
40	548	47.7	692	12	BG699266	BG699266 602678938
41	535	46.5	866	9	AU079251	AU079251 AU079251
42	530	46.1	592	13	BM023074	BM023074 id65h09.y
43	528.5	46.0	767	12	BF608560	BF608560 MY1_00150
44	528.5	46.0	939	14	BQ947443	BQ947443 AGENCOURT
45	527	45.8	404	13	BM149006	BM149006 TCAAP2E57

ALIGNMENTS

RESULT 1
BM921365
LOCUS BM921365 1014 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6626159 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752779 5', mRNA sequence.
ACCESSION BM921365
VERSION BM921365.1 GI:19371744
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12787  row: i  column: 04
            High quality sequence stop: 735.
FEATURES             Location/Qualifiers
     source          1. .1014
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5752779"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      237 a    288 c    295 g    194 t
ORIGIN

Alignment Scores:
Pred. No.:                6.16e-139      Length:          1014
Score:                    1150.00        Matches:         223
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       14             Gaps:            0

US-09-817-199B-2 (1-223) x BM921365 (1-1014)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 107

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 167

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 227

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 287

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 347

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 407

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 467

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 527

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 587

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 647

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 707

Qy        221 SerPheMet 223
              |||||||||
Db        708 TCCTTCATG 716

RESULT 2
BG283602
LOCUS       BG283602               1051 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602407930F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520191 5',
            mRNA sequence.
ACCESSION   BG283602
VERSION     BG283602.1  GI:13033710
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10417  row: k  column: 08
            High quality sequence stop: 782.
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4520191"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      275 a    287 c    314 g    175 t
ORIGIN

Alignment Scores:
Pred. No.:                3.24e-137      Length:          1051
Score:                    1137.00        Matches:         223
Percent Similarity:       99.55%         Conservative:    0
Best Local Similarity:    99.55%         Mismatches:      0
Query Match:              98.87%         Indels:          1
DB:                       12             Gaps:            0

US-09-817-199B-2 (1-223) x BG283602 (1-1051)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 62
```

```
Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 122

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 182

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 242

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 302

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 362

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 422

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 482

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 542

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 602

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGln-LysLysArgSerSerCysCy 220
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        603 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAAGAAGCGCTCCAGCTGCTG 662

Qy        220 sSerPheMet 223
              ||||||||||
Db        663 CTCCTTCATG 672

RESULT 3
AL522282
LOCUS       AL522282                887 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL522282 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 5
            prime, mRNA sequence.
ACCESSION   AL522282
VERSION     AL522282.1  GI:12785775
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 887)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .887
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB008YF08"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      203 a    256 c    257 g    168 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:                6.08e-137      Length:          887
Score:                    1134.00        Matches:         220
Percent Similarity:       99.10%         Conservative:    1
Best Local Similarity:    98.65%         Mismatches:      2
Query Match:              98.61%         Indels:          0
DB:                       9              Gaps:            0

US-09-817-199B-2 (1-223) x AL522282 (1-887)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 87

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 207

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||
Db        208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAAAGTGAAG 267

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 327

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 387

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 447

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        448 CTAGGCAACAAGGCGGATATGAGCAGCGAAARAGTGATCCGTTCCGAAGACGGAGAGACC 507

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 567

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 627

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 687

Qy        221 SerPheMet 223
              |||||||||
Db        688 TCCTTCATG 696

RESULT 4
```

```
BB598938
LOCUS       BB598938                676 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB598938 RIKEN full-length enriched, adult pancreas islet cells Mus
            musculus cDNA clone C820003E14 5', mRNA sequence.
ACCESSION   BB598938
VERSION     BB598938.2  GI:16450366
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Dec 1, 2000 this sequence version replaced gi:11507539.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES            Location/Qualifiers
     source         1. .676
                    /organism="Mus musculus"
                    /db_xref="taxon:10090"
                    /clone="C820003E14"
                    /clone_lib="RIKEN full-length enriched, adult pancreas
                    islet cells"
                    /tissue_type="pancreas"
                    /cell_type="islet cells"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                    prepared and sequenced in Mouse Genome Encyclopedia
                    Project of Genome Exploration Research Group in Riken
                    Genomic Sciences Center and Genome Science Laboratory in
                    RIKEN. Division of Experimental Animal Research in Riken
                    contributed to prepare mouse tissues. 1st strand cDNA was
                    primed with a primer [5'
                    GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                    prepared by using trehalose thermo-activated reverse
                    transcriptase and subsequently enriched for full-length by
                    cap-trapper. Second strand cDNA was prepared with the
                    primer adapter of sequence [5'
                    GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA
                    was cleaved with BamHI and XhoI. Vector: a modified
                    pBluescript KS(+) after bulk excision from Lambda FLC I.
                    Cells were provided by Tomohiro Kono, Department of Animal
                    Science, Tokyo University of Agriculture, Atsugi  City,
                    Kanagawa Prefecture, Japan, whose assistance we gratefully
                    acknowledge."
BASE COUNT      157 a    178 c    198 g    142 t      1 others
ORIGIN
```

Alignment Scores:

Pred. No.:	3.24e-127	Length:	676
Score:	1058.00	Matches:	205
Percent Similarity:	95.45%	Conservative:	5
Best Local Similarity:	93.18%	Mismatches:	10
Query Match:	92.00%	Indels:	0
DB:	10	Gaps:	0

US-09-817-199B-2 (1-223) x BB598938 (1-676)

```
Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              |||||||||||||||||||||              ||||||||||||||||||||||||||
Db         16 ATGACTGGCACACCAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCG 75

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||   ||||||:::||||||||||||||||||||||||||||||||||:::||||||
Db         76 CCCTTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTC 135

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 195

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||   ||||||||||||||||||||||||||   |||||||||
Db        196 GCCACCGTCGGCATAGACTCCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAG 255

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTAC 315

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        316 CGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACAAC 375

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 ATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGAGACGTGGTGATTATGCTT 435

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        436 CTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACA 495

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        496 CTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTG 555

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||   ||||||||||||||||||||||||||||   |||   |||
Db        556 GAGTTGGCCTTTCTGGNCATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTGAT 615

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 675
```

RESULT 5
BB633978
LOCUS BB633978 651 bp mRNA linear EST 26-OCT-2001
DEFINITION BB633978 RIKEN full-length enriched, adult male spinal cord Mus musculus cDNA clone A330094K10 5', mRNA sequence.
ACCESSION BB633978
VERSION BB633978.1 GI:16470310
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 651)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .651
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A330094K10"
/clone_lib="RIKEN full-length enriched, adult male spinal cord"
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. "
BASE COUNT 151 a 168 c 192 g 140 t
ORIGIN

Alignment Scores:

Pred. No.:	2.76e-122	Length:	651
Score:	1020.00	Matches:	198
Percent Similarity:	95.75%	Conservative:	5
Best Local Similarity:	93.40%	Mismatches:	9
Query Match:	88.70%	Indels:	0
DB:	10	Gaps:	0

US-09-817-199B-2 (1-223) x BB633978 (1-651)

```
Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||              ||||||||||||||||||||||||||||
Db         16 ATGACTGGCACACCAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCG 75

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||   ||||||:::||||||||||||||||||||||||||||||:::||||||
Db         76 CCCTTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTC 135

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 195

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        196 GCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAG 255

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTAC 315

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        316 CGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACAAC 375

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 ATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGAGACGTGGTGATTATGCTT 435

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        436 CTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACA 495

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        496 CTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTG 555

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              |||||||||||||||   ||||||||||||||||||||||||||||||   |||   |||
Db        556 GAGTTGGCCTTTCTGGGCATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTGAT 615

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSer 212
              |||||||||||||||||||||||||||||||||||||
Db        616 GAGCCCAGCTTTCAGATCCGAGACTATGTGGAGTCC 651
```

RESULT 6
BI767046
LOCUS BI767046 740 bp mRNA linear EST 25-SEP-2001
DEFINITION 603054267F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203600 5',

```
          mRNA sequence.
ACCESSION   BI767046
VERSION     BI767046.1  GI:15758624
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11510  row: f  column: 17
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5203600"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      171 a    198 c    222 g    149 t
ORIGIN

Alignment Scores:
Pred. No.:                2.27e-120      Length:          740
Score:                    1006.00        Matches:         216
Percent Similarity:       94.37%         Conservative:    2
Best Local Similarity:    93.51%         Mismatches:      5
Query Match:              87.48%         Indels:          8
DB:                       13             Gaps:            0

US-09-817-199B-2 (1-223) x BI767046 (1-740)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 60

Qy         21 ProCysSerPro-SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVa 40
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTGCAGTCCTGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGT 120

Qy         40 lGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIl 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCAT 180

Qy         60 eAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLy 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAA 240

Qy         80 sLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTy 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTA 300

Qy        100 rArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAA 360

Qy        120 nIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValVal-IleMetL 140
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        361 CATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTTGATCATGC 420

Qy        140 euLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluT 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGA 480

Qy        160 hrLeuAlaArgGluTyrGlyValProPheLeu-GluThrSerAlaLysThr-GlyMetAs 179
              |||||   |||   ||||||||||||||||| |||   |||||||||||| ||||||||
Db        481 CCTTGGGCAGGTAGTACGGTGTTCCCTTCCTGGGAGAACAGTGCCAAGACTTGGCATGAA 540

Qy        179 nValGluLeu-AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnA 199
              |||||||||| |||||||||   ||||||:::|||||||:::|||||||||||||||||||
Db        541 TGTGGAGTTAGGCTTTTCTTGGCATCGCCAGGGAACTTGAATACCGGGCCGGGCATCAGG 600

Qy        199 laAspGluProSerPheGlnIle-ArgAspTyrValGlu-SerGlnLysLysArgSerSe 218
              ||||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||
Db        601 CGGATGAACCCAGCTTCCAGATTCCGAGACTATGTAGAAGTCCCAGAAGAAGCGCTCAAG 660

Qy        218 r-CysCysSerPheMet 223
              | |||  ||||||||
Db        661 CTTGCTGGTCCTTCATG 677

RESULT 7
BB635649
LOCUS       BB635649                655 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB635649 RIKEN full-length enriched, 0 day neonate thymus Mus
            musculus cDNA clone A430077P13 5', mRNA sequence.
ACCESSION   BB635649
VERSION     BB635649.1  GI:16471692
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
```

```
           Computer-based methods for the mouse full-length cDNA
          encyclopedia: real-time sequence clustering for construction of a
          nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
           Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
          ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
          Hayashizaki,Y.
           Computational Analysis of Full-Length Mouse cDNAs Compared with
          Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
           Please visit our web site (http://genome.gsc.riken.go.jp) for
          further details.
           e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A430077P13"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     thymus"
                     /tissue_type="thymus"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT      152 a    171 c    191 g    141 t
ORIGIN

Alignment Scores:
Pred. No.:              8.4e-120        Length:         655
Score:                  1001.00         Matches:        194
Percent Similarity:     94.34%          Conservative:   6
Best Local Similarity:  91.51%          Mismatches:     12
Query Match:            87.04%          Indels:         0
DB:                     10              Gaps:           0

US-09-817-199B-2 (1-223) x BB635649 (1-655)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||          ||||||||||||||||||||||||||||
Db         18 ATGACTGGCACACCAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCG 77

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||   ||||||:::||||||||||||||||||||||||||||||||||:::|||||||
Db         78 CCCTTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTC 137

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 197

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db        198 GCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAG 257

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTAC 317

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||

Db        318 CGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACAAC 377

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGAGACGTGGTGATTATGCTT 437

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        438 CTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACA 497

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        498 CTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTG 557

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||   ||||||||||||   :::|||||||||||||   |||   |||
Db        558 GAGTTGGCCTTTCTGGGCATTGCCAAGGAAACTGAATACCGTGCAGGGAGGCAGCCTGAT 617

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSer 212
              |||||||       |||||||||||||||||||||||||
Db        618 GAGCCCCAGCTTCAGATCCGAGACTATGTGGAGTCC 653

RESULT 8
AL559085
LOCUS       AL559085                749 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL559085 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ014YB11 5
            prime, mRNA sequence.
ACCESSION   AL559085
VERSION     AL559085.1  GI:12904236
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ014YB11"
                     /clone_lib="LTI_NFL008_TC2"
                     /sex="male"
                     /tissue_type="T cells from T cell leukemia"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime  end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      157 a    213 c    228 g    129 t     22 others
ORIGIN

Alignment Scores:
Pred. No.:              7.39e-101       Length:         749
Score:                  856.50          Matches:        180
Percent Similarity:     79.22%          Conservative:   3
Best Local Similarity:  77.92%          Mismatches:     36
Query Match:            74.48%          Indels:         12
DB:                     9               Gaps:           1

US-09-817-199B-2 (1-223) x AL559085 (1-749)
```

```
Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         33 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCSAGCGCTCCCCG 92

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db         93 CCCTGCAGTCCGAGCTACGACCTCMCGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 152

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||   |||   |||   |||||||||      |||||||||   |||
Db        153 GGCAAAACATGTTTCCYGATCCMATTCMAAGACGGGGCCYTCCYGTCCGGAACCYTCATA 212

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GCCMCCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 272

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||   ||||||||||||||   ||||||||   ||||||||||
Db        273 CTGCAGATCTGGGACACCSCYGGGCAGGAACGGTTCCSAAGCGTCACCCMTGCYTATTAC 332

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 392

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        393 ATCAGGGCCTGGCTCMCTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 452

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSer-GluAspGlyGluTh 160
              ||||||||||||||||||||||||||||||||||||||||||||| :::||||||||||||
Db        453 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGCAAGACGGAGAGAC 512

Qy        160 rLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVa 180
              ||||||||||||||||||||||||||      ||||||      ||||||||||||||||
Db        513 CTTGGCCAGGGAGTACGGTGTTCCCYTCCYGGAGACCASCSCCAAGACTGGCATGAATGT 572

Qy        180 lGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAs 200
              ||||||||||||||||||||||||||||            |||   |||      |||
Db        573 GGAGTTAGCCTTTCTGGCCATCGCCAAGTGAGAGCTGGGCAGGGAAGGGAAGTGTGCGGG 632

Qy        200 pGlu----------------------------------ProSerPheGlnIleArgAspTy 209
               :::                                 |||      |||      |||
Db        633 GCAGGGCGGCACACTCCAGGAATCCAGTAGGGCCCSGCCCCYGGCCCAGCCCCTGGACAC 692

Qy        209 rValGluSerGlnLysLysArgSerSerCys 219
                      :::   ||||||   |||
Db        693 ACCTGCATTCTGCAGGCTGAGGTCCATTTGC 723
```

RESULT 9
BB619625

LOCUS	BB619625 733 bp mRNA linear EST 26-OCT-2001
DEFINITION	BB619625 RIKEN full-length enriched, 8 days embryo Mus musculus cDNA clone 5730582J11 5', mRNA sequence.
ACCESSION	BB619625
VERSION	BB619625.1 GI:16458927
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 733)
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000) wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.

FEATURES	Location/Qualifiers
source	1. .733 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="5730582J11" /clone_lib="RIKEN full-length enriched, 8 days embryo" /sex="mixed" /dev_stage="8 days embryo" /lab_host="DH10B" /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 100.0 Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT 186 a 186 c 200 g 161 t
ORIGIN

Alignment Scores:

Pred. No.:	8.3e-101	Length:	733
Score:	856.00	Matches:	187
Percent Similarity:	88.44%	Conservative:	12
Best Local Similarity:	83.11%	Mismatches:	20
Query Match:	74.43%	Indels:	7
DB:	10	Gaps:	0

US-09-817-199B-2 (1-223) x BB619625 (1-733)

```
Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              |||||||||   ||||||            ||||||||||||||||||||||||||||||
```

```
Db          3 ACTGGCACACTAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCGCCC 62

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              ||||||:::|||||||||||||||||||||||||||||||||||:::|||||||||
Db         63 TTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTCGGC 122

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 182

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        183 ACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAGCTT 242

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTACCGA 302

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
              ||||||||||||||   |||||||||||||||||||:::      ||||||||||||
Db        303 GATGCTCAAGCTTTGTTCCTGTTGTATGACATCACCAACCAATTCTTTTTTGACAACATA 362

Qy        122 ArgAlaTrpLeuThr-GluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLe 141
              |||||||||||||| :::|||||||||   ||||||||||||||||||||||||||||
Db        363 AGGGCCTGGCTTACAAAAGATTCATGAAGTTGCCCAGAGAGACGTGGTGATTATGCTTTT 422

Qy        141 uGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLe 161
              ||||   |||||||||:::|||||||||||||||||||||||||:::|||:::||||||
Db        423 AGGCAC-AAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGAAGGAAAGACACT 481

Qy        161 uAlaArgGluTyrGlyValProPheLeuGlu-ThrSerAlaLysThrGlyMetAsnValG 181
              ||||||||||||||||   ||||||:::||| |||   ||||||||||||:::|||||||
Db        482 GGCCAGGGAATATGGTGGTCCTTTCATGGAGAACCGTGGCCAAGACTGGCTTGAACGTGG 541

Qy        181 luLeuAla-PheLeuAlaIleAlaLysGlu-LeuLysTyrArgAla-GlyHisGlnAlaA 200
              |||||||| |||||||||||||||||||| |||||||||||||| |||      :::|
Db        542 AGTTGGCCCTTTCTGGCAATTGCCAAGGAACTTAAAATACCGTGCAAGGAGGCAGCTCTG 601

Qy        200 spGluProSerPheGlnIleArgAspTyrValGlu-SerGlnLysLysArgSerSerCys 219
              ||||||||   ||||||   |||:::|||||||| |||||||||   |||||||:::
Db        602 ATGAACCCCACTTTCAAAACCGAGAATATGTGGAAGTCCCAAAAAAACCGCTCCAACTGG 661

Qy        220 Cys 220
              |||
Db        662 TGC 664
```

```
RESULT 10
BM151643
LOCUS       BM151643                 522 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCBAP1E10302 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP1030, mRNA
            sequence.
ACCESSION   BM151643
VERSION     BM151643.1  GI:17175247
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCBAP1030"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamHI-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      119 a    146 c    158 g     99 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:                1.6e-100        Length:                 522
Score:                    852.00          Matches:                165
Percent Similarity:       100.00%         Conservative:           0
Best Local Similarity:    100.00%         Mismatches:             0
Query Match:              74.09%          Indels:                 0
DB:                       13              Gaps:                   0

US-09-817-199B-2 (1-223) x BM151643 (1-522)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 78

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 138

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 198

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 258

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 318

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 AGACGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 378

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 438

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 498

Qy        161 LeuAlaArgGluTyr 165
              |||||||||||||||
```

```
Db        499 TTGGCCAGGGAGTAC 513

RESULT 11
BG934396/c
LOCUS       BG934396                736 bp    mRNA    linear   EST 02-OCT-2001
DEFINITION  SK1-0698 Atlantic Salmon kidney Salmo salar cDNA clone SK1-0698
            similar to GTPase Rab37, mRNA sequence.
ACCESSION   BG934396
VERSION     BG934396.1  GI:15844224
KEYWORDS    EST.
SOURCE      Atlantic salmon.
  ORGANISM  Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Douglas,S.E., Tsoi,S.C.M., Penny,S., Melville,K., Reith,M.E. and
            Ewart,K.V.
  TITLE     Expressed Sequence Tags - A Snapshot of the Fish Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Douglas S
            Genome Sciences
            NRC Institute for Marine Biosciences
            1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada
            Tel: (902) 426-4991
            Fax: (902) 426-9413
            Email: susan.douglas@nrc.ca
            Seq primer: SK.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Salmo salar"
                     /db_xref="taxon:8030"
                     /clone="SK1-0698"
                     /clone_lib="Atlantic Salmon kidney"
                     /dev_stage="adult"
                     /note="Organ: kidney; Vector: Lambda ZAP II"
BASE COUNT      146 a    220 c    180 g    190 t
ORIGIN

Alignment Scores:
Pred. No.:              2.78e-97        Length:         736
Score:                  829.00          Matches:        155
Percent Similarity:     90.26%          Conservative:   21
Best Local Similarity:  79.49%          Mismatches:     19
Query Match:            72.09%          Indels:         0
DB:                     13              Gaps:           0

US-09-817-199B-2 (1-223) x BG934396 (1-736)

Qy         26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
              :::|||:::     |||||||||||||||||||||:::   |||||||||||||||||
Db        666 TTTGATATTGCCTTCAAGGTGATGCTGCTGGGAGACTCTGCGGTGGGGAAGACATGTGTG 607

Qy         46 LeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIle 65
              |||::::::|||||||||||||||||||||   |||   ||||||||||||||||||||
Db        606 CTGGTGCGCTTTAAAGACGGGGCCTTTTTGGGAGGCAACTTCATAGCCACCGTGGGAATA 547

Qy         66 AspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAsp 85
              ||||||||||||||||||||||||||||     :::|||||||||||||||||||||||
Db        546 GACTTTAGGAATAAAGTTGTGACGGTGGACAACGCGAAGGTCAAACTCCAGATCTGGGAT 487

Qy         86 ThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAla 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 ACTGCAGGACAGGAGAGATTCCGCAGTGTCACGCACGCCTACTACAGAGACGCACAGGCC 427

Qy        106 LeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeu 125
              ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        426 CTCCTCCTGCTCTATGACATCACCAGCAAGTCATCCTTTGACAACATCAGGGCCTGGCTG 367

Qy        126 ThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAla 145
              ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        366 ACTGAAATCCATGAGTATGCTCAGAAGGATGTGGTCATCATGTTGCTCGGCAACAAGACG 307

Qy        146 AspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyr 165
              |||||||:::   |||||||||||||:::   |||:::||||||   ||||||:::||||||
Db        306 GACATGGCCGGAGAGAGGGTCATCAAGCGTGAGGAGGGAGAGAAGCTGGCCAAGGAATAT 247

Qy        166 GlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeu 185
              ||||||||||||:::||||||||||||||||||||||:::||||||||||||||||||||
Db        246 GGAGTCCCTTTCATGGAGACCAGTGCCAAGACTGGAGTCAACGTAGAGCTGGCCTTCCTG 187

Qy        186 AlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGln 205
              |||:::|||||||||||||||:::||||||     |||   :::||||||   ||||||
Db        186 GCTGTAGCAAAGGAGTTGAAGCACAGAGCTGCCCAGCAGCCCAACGAGCCCAAGTTCCAG 127

Qy        206 IleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              |||   |||||||:::   |||:::|||:::::::|||   ||||||
Db        126 ATCCACGACTACATCGCGTCTGAGAAGCAGAAGTCTGGCTGTTGT 82

RESULT 12
BB206788
LOCUS       BB206788                676 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  BB206788 RIKEN full-length enriched, 0 day neonate thymus Mus
            musculus cDNA clone A430077P13 3', mRNA sequence.
ACCESSION   BB206788
VERSION     BB206788.2  GI:16352175
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jun 30, 2000 this sequence version replaced gi:8871741.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
```

```
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A430077P13"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     thymus"
                     /tissue_type="thymus"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT      166 a    167 c    196 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:              8.07e-97        Length:         676
Score:                  825.00          Matches:        161
Percent Similarity:     98.21%          Conservative:   4
Best Local Similarity:  95.83%          Mismatches:     3
Query Match:            71.74%          Indels:         0
DB:                     10              Gaps:           0

US-09-817-199B-2 (1-223) x BB206788 (1-676)

Qy         56 SerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAsp 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TCGGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGAT 62

Qy         76 GlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerVal 95
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GGTGCCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTG 122

Qy         96 ThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLys 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        123 ACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAG 182

Qy        116 SerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAsp 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TCCTCTTTTGACAACATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGAGAC 242

Qy        136 ValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSer 155
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        243 GTGGTGATTATGCTTCTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCT 302

Qy        156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        303 GAAGATGGAGAGACACTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAG 362

Qy        176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ACTGGCATGAACGTGGAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAATACCGTGCA 422

Qy        196 GlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLys 215
```

```
              ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GGGAGGCAGCCTGATGAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAGAAGAAG 482

Qy        216 ArgSerSerCysCysSerPheMet 223
              |||||||||||||||||||||:::
Db        483 CGCTCCAGCTGCTGCTCCTTTGTG 506

RESULT 13
BE755280
LOCUS       BE755280                553 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  209068 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE755280
VERSION     BE755280.1  GI:10169272
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 57  row: E  column: 12
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      111 a    167 c    174 g    101 t
ORIGIN

Alignment Scores:
Pred. No.:              2.16e-92        Length:         553
Score:                  790.00          Matches:        151
Percent Similarity:     97.52%          Conservative:   6
Best Local Similarity:  93.79%          Mismatches:     4
Query Match:            68.70%          Indels:         0
DB:                     12              Gaps:           0

US-09-817-199B-2 (1-223) x BE755280 (1-553)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db         71 ATGACGGGAACGCCTGGCACCGCGGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCA 130

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||  ||||||:::||||||||||||||||||||||||||||||||:::|||||||
```

```
Db        131 CCCTGCGGGCCGAGCCACGATCTCACGGGCAAGGTGATGCTTCTGGGAGACTCGGGCGTC 190

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||:::|||:::|||||||||||||||||||||||||||||||
Db        191 GGCAAAACCTGTTTCCTCATCCGATTCAGAGACGGGGCCTTCCTGTCCGGGACCTTCATA 250

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GCCACCGTCGGCATAGACTTCAGGAACAAAGTGGTGACCGTGGATGGTGTGAGAGTGAAA 310

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CTGCAGATCTGGGACACGGCAGGGCAGGAGCGGTTCCGCAGTGTGACCCATGCTTATTAC 370

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CGAGACGCCCAGGCCTTGCTCCTGCTGTACGACATCACCAACAAATCTTCCTTTGACAAC 430

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 ATCCGGGCTTGGCTCACTGAGATTCACGAGTACGCCCAGAGGGATGTGGTGATCATGCTT 490

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||:::|||:::||||||||||||||||||||||||||||||||
Db        491 CTGGGCAACAAGGCAGATGTGAGCACTGAAAGGGTGATCCGCTCAGAGGACGGAGAGATG 550

Qy        161 Leu 161
              |||
Db        551 CTG 553

RESULT 14
BM149118
LOCUS       BM149118                488 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCAAP2E6371 Pediatric acute myelogenous leukemia cell (FAB M1)
            Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP6371, mRNA
            sequence.
ACCESSION   BM149118
VERSION     BM149118.1  GI:17170448
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP6371"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project=TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      112 a    134 c    145 g     96 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                5.88e-89        Length:          488
Score:                    763.00          Matches:         152
Percent Similarity:       97.48%          Conservative:    3
Best Local Similarity:    95.60%          Mismatches:      4
Query Match:              66.35%          Indels:          1
DB:                       13              Gaps:            0

US-09-817-199B-2 (1-223) x BM149118 (1-488)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 71

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||:::   ||||||||||||||||||||||||||||||||||||||
Db         72 CCCTGCAGTCCGAGCTTTTT-CTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 130

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        131 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGCCGCCTTCCTGTCCGGAACCTTCATA 190

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 250

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        251 CTGCAGATCTGGGACACCGCTGNGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 310

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||   :::||||||||||||||||
Db        311 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAAACAATCTTCTTTCGACAAC 370

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 430

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        431 CTAGGCAACAGGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAG 487

RESULT 15
BQ687186
LOCUS       BQ687186                888 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8034867 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6206869
            5', mRNA sequence.
ACCESSION   BQ687186
VERSION     BQ687186.1  GI:21812502
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
           Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM2362  row: a  column: 14
           High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6206869"
                     /clone_lib="NIH_MGC_110"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of  California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      172 a    279 c    289 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              3.77e-88        Length:         888
Score:                  760.00          Matches:        149
Percent Similarity:     76.62%          Conservative:   28
Best Local Similarity:  64.50%          Mismatches:     30
Query Match:            66.09%          Indels:         24
DB:                     14              Gaps:           3

US-09-817-199B-2 (1-223) x BQ687186 (1-888)

Qy          4 ThrProGlyAla------ValAlaThrArgAspGlyGluAlaProGluArgSer------ 19
              ||||||   |||      :::   |||   :::|||      |||   ||||||
Db         81 ACCCCCGCTGCCTCCACGCTGCCCACCGCCAACGGGGCCCGACCGGCGCGCTCCGGGACT 140

Qy         20 --------------------ProProCysSerPro---------------------- 24
                                 ||||||      |||
Db        141 GCGCTTTCCGGCCCCGACGCGCCGCCCAACGGGCCCTTGCAGCCCGGCCGGCCCTCGCTT 200

Qy         25 --------------SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39
                           |||||||:::      |||||||||||||:::||||||:::|||
Db        201 GGCGGCGGTGTCGACTTCTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGT 260

Qy         40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59
              ||||||||||||||   |||::::::|||||||||||||||||||:::||||||||||
Db        261 GTGGGGAAGACCTGTCTGCTGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTC 320

Qy         60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79
              |||:::||||||||||||||||||||||||||:::   ||||||||||||:::|||
Db        321 ATCTCCACCGTAGGCATTGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTG 380

Qy         80 LysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AAGCTGCAGATGTGGGACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTAC 440

Qy        100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAsp 119
              ||||||||||||   ||||||||||||||||||||:::||||||||:::||||||||||
Db        441 TACCGGGATGCTCATGCTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGAC 500

Qy        120 AsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMet 139
              ||||||:::||||||||||||||||||||||||||||||||   ||||||   :::|||
Db        501 AACATCCAGGCCTGGCTGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATG 560

Qy        140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159
              ||||||||||||||   |||   :::   |||||||||:::::   ||||||||||||
Db        561 CTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAG 620

Qy        160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
                 ||||||:::|||||||||||:::||||||:::|||||||||||||||||||||:::|||
Db        621 AAGCTGGCCAAGGAGTATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAAC 680

Qy        180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla 199
              |||:::|||||||||   |||||||||||||||||:::|||   |||:::
Db        681 GTGGACTTGGCCTTCACAGCCATAGCAAAGGAGGTGAAGCAGCGCTCCCTGAAGGCTCCC 740

Qy        200 AspGluProSerPheGlnIleArgAspTyrVal 210
                 ||||||   |||::::::   ||||||||
Db        741 AGCGAGCCGCGCTTCCGGCTGCATGATTACGTT 773


Search completed: June 23, 2003, 20:56:39
Job time : 1388 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 15:56:08 ; Search time 228 Seconds
(without alignments)
2202.612 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09817199/runat_18062003_145101_785/app_query.fasta_1.391
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199_@CGN_1_1_263_@runat_18062003_145101_785 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	1106	22	AAI59066	Human polynucleoti
2	1150	100.0	2323	22	AAI60852	Human polynucleoti
3	1145	99.6	1109	22	AAH75184	Nucleotide sequenc
4	1144	99.5	2623	22	AAS27036	cDNA encoding nove
5	1140	99.1	875	24	ABN83696	Human Ras protein
6	1133	98.5	1316	22	AAS27458	cDNA encoding nove
7	1034	89.9	843	20	AAX86720	DNA sequence encod
8	969	84.3	576	24	ABA02774	Human degranulatio
9	942	81.9	576	24	ABA02773	Mouse degranulatio
10	770	67.0	475	23	AAS81431	DNA encoding novel
11	731	63.6	1340	19	AAV65197	Human RAB protein,
12	697	60.6	2001	23	ABL29523	Drosophila melanog
13	535	46.5	964	22	AAS27040	cDNA encoding nove
14	516	44.9	911	21	AAC38429	Arabidopsis thalia
15	515.5	44.8	1129	21	AAC44482	Zea mays DNA fragm
16	515	44.8	1203	21	AAC43685	Zea mays DNA fragm
17	514.5	44.7	2411	24	ABQ54410	Human ovarian anti
18	513.5	44.7	1986	22	ABA09160	Human rab8 homolog
19	511.5	44.5	2497	22	AAS60878	Human cancer agent
20	511.5	44.5	2497	22	AAS60884	Human cancer agent
21	511.5	44.5	2497	22	AAS60893	Human cancer agent
22	511.5	44.5	2497	22	AAS60895	Human cancer agent
23	511.5	44.5	2497	23	ABV25781	Human prostate exp
24	511.5	44.5	2497	23	ABV30037	Human prostate exp
25	508.5	44.2	1540	23	ABL27707	Drosophila melanog
26	507	44.1	777	21	AAC42684	Arabidopsis thalia
27	507	44.1	881	24	ABN98761	Arabidopsis thalia
28	507	44.1	1023	21	AAC51491	Arabidopsis thalia
29	507	44.1	1025	21	AAC34347	Arabidopsis thalia
30	507	44.1	3077	22	AAH13912	Human cDNA sequenc
31	506.5	44.0	956	21	AAA96887	Nucleotide sequenc
32	506	44.0	674	23	AAS71453	DNA encoding novel
33	506	44.0	1537	21	AAF16196	Human prostate can
34	506	44.0	3533	22	AAH17889	Human cDNA sequenc
35	505.5	44.0	716	21	AAA40104	Human Rab10 cDNA.
36	505.5	44.0	861	21	AAA40108	Human Rab10 cDNA #
37	503	43.7	959	21	AAC33987	Arabidopsis thalia
38	502	43.7	1101	21	AAC34080	Arabidopsis thalia
39	501	43.6	866	22	AAH04301	Human cDNA clone (
40	501	43.6	1161	22	AAH75182	Nucleotide sequenc
41	498	43.3	888	21	AAA40109	Canine Rab10 cDNA.
42	497.5	43.3	932	21	AAC35200	Arabidopsis thalia
43	490	42.6	2247	23	ABL29661	Drosophila melanog
44	488	42.4	705	21	AAC49208	Arabidopsis thalia
45	483	42.0	666	21	AAC42764	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAI59066
ID AAI59066 standard; cDNA; 1106 BP.
XX
AC AAI59066;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1269.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens

XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39910.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1269; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1106 BP; 260 A; 307 C; 321 G; 218 T; 0 other;

Alignment Scores:

Pred. No.:	2.36e-128	Length:	1106
Score:	1150.00	Matches:	223
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-817-199B-2 (1-223) x AAI59066 (1-1106)

```
Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 87

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 207

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 267

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 327

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 387

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 447

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 507

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 567

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 627

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 687

Qy        221 SerPheMet 223
              ||||||||
Db        688 TCCTTCATG 696
```

RESULT 2
AAI60852
ID AAI60852 standard; cDNA; 2323 BP.
XX
AC AAI60852;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4841.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

```
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41696.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4841; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Alignment Scores:
Pred. No.:                6.74e-128      Length:          2323
Score:                    1150.00        Matches:         223
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       22             Gaps:            0

US-09-817-199B-2 (1-223) x AAI60852 (1-2323)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 87

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 207

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 267

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 327

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 387

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 447

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 507

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 567

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 627

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 687

Qy        221 SerPheMet 223
              ||||||||
Db        688 TCCTTCATG 696

RESULT 3
AAH75184
ID   AAH75184 standard; DNA; 1109 BP.
XX
AC   AAH75184;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Nucleotide sequence of human 32712 G-protein.
XX
KW   Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW   congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW   adult respiratory distress syndrome; Goodpasture's syndrome;
KW   chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
KW   liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
KW   cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
KW   brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
KW   acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
KW   chronic bacterial meningoencephalitis; multiple sclerosis;
KW   amyotropic lateral sclerosis; stroke; Huntington's disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             124..699
FT                   /*tag= a
FT                   /product= "G-protein"
XX
PN   WO200164887-A2.
XX
PD   07-SEP-2001.
XX
PF   27-FEB-2001; 2001WO-US06292.
XX
PR   29-FEB-2000; 2000US-0185606.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers RA;
XX
DR   WPI; 2001-550182/61.
DR   P-PSDB; AAG67156.
XX
PT   Novel human small G-protein polypeptides and polynucleotides for
PT   treating lung disorders, liver disorders and brain disorders  -
XX
PS   Claim 2; Fig 26; 151pp; English.
XX
CC   The present sequence encodes a human G-protein. The specification
CC   describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
CC   G-protein polypeptides and polynucleotides are useful as a target for
CC   diagnosis and treatment of  G-protein mediated or related disorders,
CC   and for identifying agonists and antagonists for diagnosis and
CC   treatment. They are useful for treating disorders of lung (e.g.
CC   congenital anomalies, pulmonary congestion, oedema, adult respiratory
CC   distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
CC   asthma, Goodpasture's syndrome and pulmonary hypertension), liver
CC   (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
CC   Wilson's disease, autoimmune hepatitis and hepatic failure), and
CC   brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
```

```
CC   meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
CC   bacterial meningoencephalitis, multiple sclerosis, amyotropic lateral
CC   sclerosis, stroke and Huntington's disease).
XX
SQ   Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

Alignment Scores:
Pred. No.:                9.45e-128      Length:          1109
Score:                    1145.00        Matches:         222
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              99.57%         Indels:          0
DB:                       22             Gaps:            0

US-09-817-199B-2 (1-223) x AAH75184 (1-1109)

Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 90

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 150

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 210

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 270

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 330

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 390

Qy        122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 450

Qy        142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 510

Qy        162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 570

Qy        182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 630

Qy        202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 690

Qy        222 PheMet 223
              ||||||
Db        691 TTCATG 696

RESULT 4
AAS27036
ID   AAS27036 standard; cDNA; 2623 BP.
XX
AC   AAS27036;
XX
DT   07-NOV-2001  (first entry)
XX
DE   cDNA encoding novel signal transduction pathway protein, Seq ID 71.
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW   antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis C; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01312.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
```

PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR P-PSDB; AAU17119.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 71; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Alignment Scores:

Pred. No.:	4.22e-127	Length:	2623
Score:	1144.00	Matches:	222
Percent Similarity:	99.55%	Conservative:	0
Best Local Similarity:	99.55%	Mismatches:	1
Query Match:	99.48%	Indels:	0
DB:	22	Gaps:	0

US-09-817-199B-2 (1-223) x AAS27036 (1-2623)

```
Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 71

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 131

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 191

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        192 GCCACCGTCGGCATAGACTTCAGCAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 251
```

```
Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 311

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 371

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 431

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 491

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 551

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 611

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 671

Qy        221 SerPheMet 223
              |||||||||
Db        672 TCCTTCATG 680
```

```
RESULT 5
ABN83696
ID   ABN83696 standard; cDNA; 875 BP.
XX
AC   ABN83696;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Ras protein 3 (RASP-3) cDNA.
XX
KW   Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;
KW   immunosuppressive; antiinflammatory; signal transduction; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             93..668
FT                   /*tag=  a
FT                   /product= "RASP-3"
XX
PN   US6391580-B1.
XX
PD   21-MAY-2002.
XX
PF   08-MAY-1998;   98US-0075454.
XX
PR   12-DEC-1996;   96US-0766551.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Tang YT,  Lal P,  Guegler KJ,  Corley NC,  Patterson C;
PI   Batra S,  Baughn MR;
XX
DR   WPI; 2002-498774/53.
DR   P-PSDB; ABB76426.
XX
PT   New isolated polynucleotides encoding Ras proteins designated RASP-1
PT   and RASP-4, for diagnosing, preventing and treating disorders
PT   associated with cell proliferation, particularly cancer and immune
PT   disorders -
XX
PS   Example; Column 57-60; 34pp; English.
XX
CC   The present sequence is that of a cDNA clone encoding novel human
CC   Ras protein 3 (RASP-3, see ABB76426).  Nucleic acids encoding
CC   RASP-3 were initially identified in Incyte Clone 1528559 from a
CC   mononuclear cell cDNA library (UCMCL5T01) using a computer search
CC   for amino acid sequence alignments.  The present consensus
CC   sequence was derived from overlapping and/or extended nucleic
CC   acids in Incyte clone 1528559 and shot-gun sequence SAEA03135,
CC   SAEC10396 and SAEC10855.  RASP-3 shows homology to rat Rab26.
CC   Northern analysis showed expression of RASP-3 in haematopoietic
CC   and immunological cDNA libraries, all associated with inflammation
CC   and the immune response.  The invention provides 7 novel human Ras
CC   proteins (RASP-1 to -7) and polynucleotides, expression vectors,
CC   host cells, antibodies, agonists and antagonists.  It also provides
CC   methods for diagnosing, treating or preventing disorders associated
CC   with RASP expression, especially cancer and immune disorders.  A
CC   fragment of the present sequence, from about nucleotide 92 to about
CC   nucleotide 153, is useful as a hybridisation probe.
XX
SQ   Sequence 875 BP; 203 A; 247 C; 260 G; 165 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              2.7e-127        Length:         875
Score:                  1140.00         Matches:        221
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            99.13%          Indels:         0
DB:                     24              Gaps:           0

US-09-817-199B-2 (1-223) x ABN83696 (1-875)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProProCys 22
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGC 62

Qy         23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAA 122

Qy         43 ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACC 182

Qy         63 ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAG 242

Qy         83 IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 ATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGAT 302

Qy        103 AlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGG 362

Qy        123 AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGC 422

Qy        143 AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCC 482

Qy        163 ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTA 542

Qy        183 AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCC 602
```

```
Qy        203 SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTC 662

Qy        223 Met 223
              |||
Db        663 ATG 665

RESULT 6
AAS27458
ID   AAS27458 standard; cDNA; 1316 BP.
XX
AC   AAS27458;
XX
DT   07-NOV-2001  (first entry)
XX
DE   cDNA encoding novel signal transduction pathway protein, Seq ID 493.
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW   antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis C; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01312.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
```

```
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465460/50.
DR   P-PSDB; AAU17541.
XX
PT   Novel polypeptides useful for diagnosing, treating, preventing and/or
PT   prognosing disorders related to the proteins, including cancers, immune
PT   disorders and neuronal disorders  -
XX
PS   Claim 1; SEQ ID No 493; 880pp; English.
XX
CC   The invention relates to novel isolated polypeptides (I), and
CC   polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC   diagnosing, preventing and treating diseases including immune system
CC   disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC   disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC   transplant rejections and graft versus host disease, infectious diseases
CC   (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC   other blood-related disorders (sickle cell anaemia), myeloproliferative
CC   disorders, primary haematopoietic disorders, hyperproliferative
CC   disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC   disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC   abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC   disorders (e.g. glomerulonephritis), cardiovascular disorders
CC   (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC   wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC   Addison's disease), reproductive system disorders, gastrointestinal
CC   disorder (inflammatory disorders), liver disorders (cirrhosis),
CC   as stimulators of B-cell responsiveness to pathogens, activators of
CC   T-cells, to induce higher affinity antibodies, and as a means to induce
CC   tumour proliferation in pathologies e.g. acquired immune deficiency
CC   syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC   pathway protein coding sequences and PCR primers of the invention.
XX

Alignment Scores:
Pred. No.:              3.34e-126     Length:          1316
Score:                  1133.00       Matches:         220
Percent Similarity:     99.10%        Conservative:    0
Best Local Similarity:  99.10%        Mismatches:      2
Query Match:            98.52%        Indels:          0
DB:                     22            Gaps:            0

US-09-817-199B-2 (1-223) x AAS27458 (1-1316)

Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 60

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 120

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 180

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGTCGGCATAGACTTCAGSAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 240

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 300

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 360

Qy        122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 420

Qy        142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 480

Qy        162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAGGGAGTACGGKGTTCCCTTYCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 540

Qy        182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 600

Qy        202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        601 CCCAGCTTCCAGATCCGAGACTATGTAKAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 660

Qy        222 PheMet 223
              ||||||
Db        661 TTCATG 666

RESULT 7
AAX86720
ID   AAX86720 standard; DNA; 843 BP.
XX
AC   AAX86720;
XX
DT   27-OCT-1999  (first entry)
XX
DE   DNA sequence encoding an exocytotic protein designated Exo2.
XX
KW   Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
KW   inflammation; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             124..699
FT                   /*tag=  a
FT                   /transl_except= (pos: 136..138, aa: Xaa)
FT                   /transl_except= (pos: 202..204, aa: Xaa)
FT                   /note= "Xaa is an unspecified amino acid"
XX
```

```
PN   WO9942586-A2.
XX
PD   26-AUG-1999.
XX
PF   23-FEB-1999;   99WO-US03909.
XX
PR   26-MAY-1998;   98US-0086650.
PR   23-FEB-1998;   98US-0075534.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Fisher J,  Huang B,  Lorens J,  Luo Y;
PI   Shen M;
XX
DR   WPI; 1999-518605/43.
DR   P-PSDB; AAY30132.
XX
PT   New exocytotic proteins useful for diagnosis and treatment of
PT   exocytosis-mediated conditions and in drug screening
XX
PS   Claim 6; Fig 6; 53pp; English.
XX
CC   The present sequence encodes an exocytotic protein designated Exo2.
CC   The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC   secretory vesicles with the cellular plasma membrane) and can be
CC   administered therapeutically to treat or prevent exocytosis-mediated
CC   disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC   undesirable release of compounds via exocytosis e.g. inflammation
CC   mediated by the release compounds such as histamine. Exo2 is useful
CC   diagnostically and to produce antibodies useful to purify the proteins
CC   and therapeutically to reduce or eliminate the biological activity of
CC   the protein. Exo2 nucleic acids can be used therapeutically to increase
CC   Exo2 activity in cells by known gene therapy techniques. They can also
CC   be used to produce probes or primers to isolate Exo2 proteins from other
CC   organisms, especially humans. The nucleic acids, host cells and proteins
CC   are useful in screening assays to identify binding agents, especially
CC   drug screening assays to identify agonists and antagonists useful
CC   therapeutically to enhance or reduce Exo2 activity.
XX
SQ   Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

Alignment Scores:
Pred. No.:              1.43e-114       Length:         843
Score:                  1034.00         Matches:        205
Percent Similarity:     94.22%          Conservative:   7
Best Local Similarity:  91.11%          Mismatches:     11
Query Match:            89.91%          Indels:         2
DB:                     20              Gaps:           0

US-09-817-199B-2 (1-223) x AAX86720 (1-843)

Qy          1 MetThrGlyThrProGlyAla--ValAlaThrArgAspGlyGluAlaProGluArgSerP 20
              |||   |||||||||||||||  :::         |||||||||||||||||||||||||
Db         26 ATGAATGGCACACCAGGAGCTTGCTACCGCTTGGGGATGGCGAGGCCCCTGAGCGCTCCC 85

Qy         20 roProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyV 40
              |||||   ||||||:::||||||   |||||||||||||||||||||||   :::||||
Db         86 CGCCCTTCAGCCCGAACTACGATNTCACCGGCAAGGTGATGCTTCTTGGAGANTCGGGCG 145

Qy         40 alGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        146 TCGGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTNA 205

Qy         60 leAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValL 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db        206 TAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTTCCAGGGTGA 265

Qy         80 ysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrT 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 AGCTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATT 325

Qy        100 yrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspA 120
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        326 ACCGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACA 385

Qy        120 snIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetL 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ACATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATTATGC 445

Qy        140 euLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluT 160
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        446 TTCTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGA 505

Qy        160 hrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnV 180
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        506 CACTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACG 565

Qy        180 alGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaA 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||   |
Db        566 TGGAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTG 625

Qy        200 spGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysC 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 ATGAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAGAAGAAGCGCTCCAGCTGCT 685

Qy        220 ysSerPheMet 223
              |||||||:::
Db        686 GCTCCTTTGTG 696

RESULT 8
ABA02774
ID   ABA02774 standard; DNA; 576 BP.
XX
AC   ABA02774;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human degranulation regulator encoding DNA SEQ ID NO 3.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..576
FT                   /*tag= a
FT                   /product= "degranulation regulator"
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
DR   P-PSDB; AAM52190.
XX
PT   Mast cell degranulation controller for treatment of allergies -
XX
PS   Claim 11; Page 64-65; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;

Alignment Scores:
Pred. No.:              5.49e-107       Length:         576
```

```
Score:                969.00        Matches:       189
Percent Similarity:   100.00%       Conservative:  2
Best Local Similarity: 98.95%       Mismatches:    0
Query Match:          84.26%        Indels:        0
DB:                   24            Gaps:          0

US-09-817-199B-2 (1-223) x ABA02774 (1-576)

Qy         33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTTGGAGACTCGGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60

Qy         53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy         73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy         93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIle 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        213 GlnLysLysArgSerSerCysCysSerPheMet 223
              ||||||||||||||||||||||||||||||:::
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTG 573

RESULT 9
ABA02773
ID   ABA02773 standard; DNA; 576 BP.
XX
AC   ABA02773;
XX
DT   07-FEB-2002 (first entry)
XX
DE   Mouse degranulation regulator encoding DNA SEQ ID NO 2.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..576
FT                   /*tag= a
FT                   /transl_except= (pos:247..249,aa:Lys)
FT                   /product= "degranulation regulator"
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T, Ido M;
XX
DR   WPI; 2002-041335/05.
DR   P-PSDB; AAM52189.
XX
PT   Mast cell degranulation controller for treatment of allergies -
XX
PS   Claim 10; Page 63-64; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Alignment Scores:
Pred. No.:            9.69e-104     Length:        576
Score:                942.00        Matches:       183
Percent Similarity:   98.43%        Conservative:  5
Best Local Similarity: 95.81%       Mismatches:    3
Query Match:          81.91%        Indels:        0
DB:                   24            Gaps:          0

US-09-817-199B-2 (1-223) x ABA02773 (1-576)

Qy         33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTTGGAGACTCGGGCGTCGGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGG 60

Qy         53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTG 120

Qy         73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAGTGGATGGTGCCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTT 180

Qy         93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIle 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATC 240

Qy        113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACCAGTCCTCTTTTGACAACATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCC 300

Qy        133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db        301 CAGAGAGACGTGGTGATTATGCTTCTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTG 360

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        361 ATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACC 420

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGCCAAGACTGGCATGAACGTGGAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAA 480

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
              ||||||||||||   |||   ||||||||||||||||||||||||||||||||||||||
Db        481 TACCGTGCAGGGAGGCAGCCTGATGAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCC 540

Qy        213 GlnLysLysArgSerSerCysCysSerPheMet 223
              ||||||||||||||||||||||||||||||:::
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTG 573

RESULT 10
AAS81431
```

```
ID   AAS81431 standard; cDNA; 475 BP.
XX
AC   AAS81431;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #17235.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17244.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 17235; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;

Alignment Scores:
Pred. No.:              3.58e-83      Length:         475
Score:                  770.00        Matches:        148
Percent Similarity:     100.00%       Conservative:   1
Best Local Similarity:  99.33%        Mismatches:     0
Query Match:            66.96%        Indels:         0
DB:                     23            Gaps:           0

US-09-817-199B-2 (1-223) x AAS81431 (1-475)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 87

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 147

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 207

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 267

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 327

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 387

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        388 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAAGGACGTGGTGATCATGCTG 447

Qy        141 LeuGlyAsnLysAlaAspMetSerSer 149
              |||||||||||||||||||||||||
Db        448 CTAGGCAACAAGGCGGATATGAGCAGC 474

RESULT 11
AAV65197
ID   AAV65197 standard; DNA; 1340 BP.
XX
AC   AAV65197;
XX
DT   17-DEC-1998  (first entry)
XX
DE   Human RAB protein, SRAB, coding sequence.
XX
KW   Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW   Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW   viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW   diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW   hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;
KW   polycystic renal disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..615
FT                   /*tag= a
FT                   /product= "SRAB"
XX
PN   WO9842839-A1.
XX
PD   01-OCT-1998.
XX
PF   25-MAR-1998;   98WO-US05871.
XX
PR   26-MAR-1997;   97US-0824873.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Guegler KJ,  Hillman JL;
XX
DR   WPI; 1998-542282/46.
DR   P-PSDB; AAW80747.
XX
PT   New human Rab protein, SRAB - useful for treating disorders
PT   associated with SRAB expression, including vesicle trafficking,
PT   viral infection, and cancer
XX
PS   Claim 5; Fig 1; 31pp; English.
```

```
XX
CC   This sequence encodes the human RAB protein, SRAB, of the invention. The
CC   SRAB protein can be used in compositions useful in the diagnosis,
CC   prevention, or treatment of disorders associated with vesicle trafficking
CC   (including epilepsy, Huntington's disease, Parkinson's disease and
CC   schizophrenia), cancer, or viral infection. Host cells containing the DNA
CC   sequence can be used to produce SRAB recombinantly. The viral infections
CC   that can be treated using SRAB include those caused by retroviruses
CC   particularly HIV and HTLV, hepadnaviruses, particularly hepatitis C,
CC   hantaviruses, herpesviruses, and arboviruses. The cancers that can be
CC   treated include adenocarcinoma, leukaemia, lymphoma, melanoma and
CC   sarcoma, particularly cancers of the endocrine, gastrointestinal and
CC   nervous systems and cancers of the adrenal gland, brain, breast, colon,
CC   oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,
CC   prostate, salivary gland, stomach, thyroid, and uterus. The protein can
CC   also be used to treat conditions such as autoimmune sialosis,
CC   choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,
CC   hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme
CC   deficiency, pancreatitis, and polycystic renal disease.
XX
SQ   Sequence 1340 BP; 276 A; 387 C; 403 G; 274 T; 0 other;

Alignment Scores:
Pred. No.:              7.63e-78      Length:          1340
Score:                  731.00        Matches:         141
Percent Similarity:     86.73%        Conservative:    29
Best Local Similarity:  71.94%        Mismatches:      25
Query Match:            63.57%        Indels:          2
DB:                     19            Gaps:            0

US-09-817-199B-2 (1-223) x AAV65197 (1-1340)

Qy         26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
              ||||||::      ||||||||||:::|||||:::|||||||||||||||||||
Db         22 TACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTG 81

Qy         46 Leu-IleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIl 65
              ||| ::::::||||||||||||||||||||||:::||||||||||||:::||||||   ||
Db         82 CTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGC-AT 140

Qy         65 eAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAs 85
              ||||||||||||||||||:::   ||||||||||||:::||||||||||||:::|||||
Db        141 TGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGA 200

Qy         85 pThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAl 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        201 CACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGC 260

Qy        105 aLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLe 125
              ||||||||||||||||||:::|||||||||:::|||||||||||||||||:::||||||||
Db        261 TCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCT 320

Qy        125 uThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAl 145
              ||||||||||||||||||||||||   ||||||   :::||||||||||||||||||||
Db        321 GACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGT 380

Qy        145 aAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTy 165
              .|||   :::   ||||||||||::::::   ||||||||||||   ||||||:::||||||
Db        381 GGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTA 440

Qy        165 rGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLe 185
              ||||:::||||||:::||||||||||||||||||||||:::|||||||:::||||||||||
Db        441 TGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCAC 500

Qy        185 uAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGl 205
              ||||||||||||||||||||   |||:::                ||||||   |||::
Db        501 AGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCG 560

Qy        205 nIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ::::   ||||||||:::   :::   :::   :::||||||||
Db        561 GCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606

RESULT 12

ABL29523
ID   ABL29523 standard; DNA; 2001 BP.
XX
AC   ABL29523;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 40042.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 40042; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2001 BP; 455 A; 616 C; 573 G; 357 T; 0 other;

Alignment Scores:
Pred. No.:              1.65e-73      Length:          2001
Score:                  697.00        Matches:         137
Percent Similarity:     74.24%        Conservative:    33
Best Local Similarity:  59.83%        Mismatches:      49
Query Match:            60.61%        Indels:          10
DB:                     23            Gaps:            3

US-09-817-199B-2 (1-223) x ABL29523 (1-2001)

Qy          3 GlyThrProGlyAlaValAlaThr---------ArgAspGlyGluAla---------- 15
              |||:::   ||||||   ||||||               |||   |||
Db       1297 GGAAGCGAGGGAGCCAGCGCAACGCTCTGCAAGAACGCTGGACGGGCACTCATCCGCATG 1356

Qy         16 ------ProGluArgSerProProCysSerProSerTyrAspLeuThrGlyLysValMet 33
                    ::::::|||               :::|||:::   |||||||||::
Db       1357 ATTTCATCGAGCAAGGCTCCGGAGGAGGAAGACGAATTCGATATCATGGGCAAGGTCATA 1416

Qy         34 LeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAla 53
              :::||||||||||:::||||||||||||||      ||||||:::|||:::||||||
Db       1417 ATGCTGGGCGATTCGGGAGTGGGCAAGACCTCACTTTTGATCCGTTTCCGCGATGGTCGC 1476

Qy         54 PheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThr 73
              ::::::         |||:::::::||||||||||||||||||||||||||||||||
```

```
Db       1477 TATGTGCCGAGCTATTTCCTCAGCACAGTTGGCATTGATTTTAGGAACAAAGTGGTGGTC 1536

Qy         74 ValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArg 93

Db       1537 GTCGATGGAACGCGCGTCAAGCTGCAAATCTGGGACACAGCTGGTCAGGAGCGATTCCGG 1596

Qy         94 SerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThr 113

Db       1597 AGCGTTACCCACGCCTATTATCGCGACGCGCACGCTCTACTGCTGCTGTACGACGTGACC 1656

Qy        114 AsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGln 133

Db       1657 AACAAGACCACCTATGACAACATTCGCGCCTGGCTGGGCGAGATCCGGGAGTACGCGCAG 1716

Qy        134 ArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSer---SerGluArgVal 152

Db       1717 GAGGACGTGGTCATCGTTTTAATAGGCAACAAGGCCGACTGCAGCGGCAGCGAGCGGCAG 1776

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172

Db       1777 GTGAAGCGGGAGGATGGGGAGCGTTTGGGGCGGGAGCACAACGTGCCCTTCATGGAGACC 1836

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192

Db       1837 TCGGCCAAGACGGGACTCAATGTGGAGCTGTCCTTCACAGCGGTGGCCAGGCAACTAAAG 1896

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212

Db       1897 AGTCGCGGCTACGAGCACGGCGATGATGGAAAGTTCAATGTGCATGATTTTGTGCGTGAC 1956

Qy        213 GlnLysLysArgSerSerCysCysSer 221

Db       1957 AATACAAAGGCGCGCTCGGTTTGCGCC 1983
```

RESULT 13
AAS27040
ID AAS27040 standard; cDNA; 964 BP.
XX
AC AAS27040;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 75.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR P-PSDB; AAU17123.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 75; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Alignment Scores:

Pred. No.:	1.79e-54	Length:	964
Score:	535.00	Matches:	98
Percent Similarity:	85.40%	Conservative:	19
Best Local Similarity:	71.53%	Mismatches:	20
Query Match:	46.52%	Indels:	0
DB:	22	Gaps:	0

US-09-817-199B-2 (1-223) x AAS27040 (1-964)

```
Qy         84 TrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAla 103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TGGGACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCT 62

Qy        104 GlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAla 123
              ||||||||||||||||||||:::||||||||:::|||||||||||||||||:::|||
Db         63 CATGCTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCC 122

Qy        124 TrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsn 143
              |||||||||||||||||||||||||||||   ||||||   :::|||||||||||||||
Db        123 TGGCTGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAAC 182

Qy        144 LysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArg 163
              |||   |||   :::   |||||||||::::::   ||||||||||   ||||||:::
Db        183 AAGGTGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAG 242

Qy        164 GluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAla 183
              |||||||||:::|||||||:::|||||||||||||||||||||:::|||||||:::||||||
Db        243 GAGTATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCC 302

Qy        184 PheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSer 203
              |||   ||||||||||||||||||||   |||:::                 ||||||
Db        303 TTCACAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGC 362

Qy        204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              |||::::::   |||||||||:::   :::   :::   :::||||||||||
Db        363 TTCCGGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 413
```

RESULT 14
AAC38429
ID AAC38429 standard; DNA; 911 BP.
XX
AC AAC38429;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 20938.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

```
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

Alignment Scores:
Pred. No.:              3.18e-52        Length:         911
Score:                  516.00          Matches:        107
Percent Similarity:     66.52%          Conservative:   48
Best Local Similarity:  45.92%          Mismatches:     58
Query Match:            44.87%          Indels:         20
DB:                     21              Gaps:           6

US-09-817-199B-2 (1-223) x AAC38429 (1-911)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              :::|||   :::|||   ||||||                  ||||||   |||:::
Db         28 CTAACATCTTCTCCGATGGCGGTT-----------------GCGCCGGCAAGAGCT--- 66

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
                        ||||||         |||:::::::|||:::|||||||:::||||||
Db         67 ------CGTTCAGACTATGATTATCTCATCAAGCTTCTCCTCATCGGTGATAGCGGTGTG 120

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||:::|||   |||:::::::|||   |||       |||   :::   :::||||||
Db        121 GGAAAAAGTTGCTTGTTACTGCGTTTTTCGGATGATACTTTCACCACG---AGTTTCATT 177

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
                 |||:::|||||||||||||:::   :::   |||   :::|||||||   |||:::|||
Db        178 ACTACCATTGGAATCGACTTCAAGATAAGAACAGTCGAACTTGATGGGAAGCGTATCAAA 237

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||:::::::|||   |||||||||
Db        238 TTGCAGATATGGGACACTGCTGGACAAGAACGTTTTAGAACTATAACCACAGCATATTAC 297

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              |||   |||      :::|||||||:::||||||:::|||:::::::||||||||||:::|||
Db        298 AGAGGAGCGATGGGTATATTACTCGTCTATGATGTAACAGATGAGTCATCATTCAACAAC 357

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||   |||:::      |||   ::::::|||         |||      :::|||
Db        358 ATTAGGAACTGGATGAAAAACATTGAGCAGCATGCGTCCGATAGTGTCAACAAGATATTG 417

Qy        141 LeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSerGluAspGlyGlu 159
              :::||||||||||||||||||      |||:::|||   :::   :::      |||:::
Db        418 GTCGGTAACAAAGCCGACATGGACGAAAGCAAAAGGGCTGTCCCAACATCAAAAGGACAA 477

Qy        160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
                 ||||||   ||||||||||:::   |||   ||||||||||||||||||      |||
Db        478 GCTCTTGCTGATGAGTATGGAATCAAATTCTTTGAGACGAGTGCAAAAACAAACCAGAAT 537

Qy        180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg---------AlaGly 196
              ||||||      ||||||:::||||||||:::::::|||   |||         :::
Db        538 GTCGAGCAGGTTTTCCTTTCTATCGCTAAAGACATAAAACAAAGACTCACAGAAAGCGAT 597

Qy        197 HisGlnAlaAspGluProSerPheGlnIle-----------------ArgAspTyrVal 210
                 :::|||:::            :::|||                 :::
Db        598 ACAAAAGCCGAGCCCCAAGGGATCAAGATCACTAAACAAGATGCTAACAAAGCCTCATCG 657

Qy        211 GluSerGlnLysLysArgSerSerCysCysSerPheMet 223
                 |||      ::::::|||:::||||||||:::::
Db        658 TCTTCTACAAATGAGAAATCAGCTTGCTGCAGTTATGTT 696

RESULT 15
AAC44482
ID   AAC44482 standard; DNA; 1129 BP.
XX
AC   AAC44482;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Zea mays DNA fragment SEQ ID NO: 42986.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic;
KW   pathway; promoter; termination sequence; corn; ss.
XX
OS   Zea mays subsp. mays.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	4.95e-52	Length:	1129
Score:	515.50	Matches:	104
Percent Similarity:	66.38%	Conservative:	50
Best Local Similarity:	44.83%	Mismatches:	61
Query Match:	44.83%	Indels:	17
DB:	21	Gaps:	6

US-09-817-199B-2 (1-223) x AAC44482 (1-1129)

```
Qy          5 ProGlyAlaValAlaThrArgAspGlyGluAlaProGluArg---------------- 18
              ||||||       :::|||   |||        |||:::
Db        127 CCGGGACCGGCGGGGTCGCGCGCCGGGTAGGAGGAGGAGAAGAGGAAGGGCGGTATGGCG 186

Qy         19 SerProPro------CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGly 36
              :::||||||                ||||||         |||:::::|||:::|||
Db        187 GCGCCGCCGGCGAGGGCCCGGGCCGACTACGACTACCTTATCAAGCTGCTCCTCATTGGA 246

Qy         37 AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer 56
              |||:::|||||||||||||:::|||   |||::::::|||   ||||||:::|||   :::
Db        247 GACAGCGGTGTTGGCAAGAGTTGCCTCCTGTTGCGGTTCTCTGATGGTTCCTTCACTACA 306

Qy         57 GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGly 76
              :::||||||   |||:::|||||||||||||:::   :::   :::   :::|||||||
Db        307 ---AGCTTTATTACCACAATTGGTATTGACTTTAAGATACGAACGATAGAATTGGATGGC 363

Qy         77 ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr 96
              |||:::|||||||||||||||||||||||||||||||||||||||||||:::::|||
Db        364 AAGCGTATAAAGCTACAGATTTGGGATACGGCGGGCCAAGAACGCTTCCGGACTATTACC 423

Qy         97 HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer 116
              ||||||||||||   |||      :::||||||:::||||||::|||::::::|||
Db        424 ACCGCGTACTACCGAGGAGCTATGGGCATCTTGCTGGTTTATGATGTCACCGATGAATCT 483

Qy        117 SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal 136
              ||||||:::|||||||||   |||:::      |||   :::::::|||     :::|||
Db        484 TCTTTTAACAACATCCGAAACTGGATTCGGAACATTGAACAACACGCCTCTGATAATGTC 543

Qy        137 ValIleMetLeuLeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSer 155
              :::|||:::|||||||||||||||||||      |||:::|||   :::   :::
Db        544 AACAAGATTTTGGTTGGCAACAAGGCCGATATGGACGAAAGCAAGAGGGCTGTACCTACT 603

Qy        156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175
              |||:::   ||||||   |||||||||:::   |||   ||||||||||||||||
Db        604 GCAAAAGGACAAGCACTTGCTGATGAGTACGGGATCAAGTTTTTTGAAACGAGTGCGAAG 663

Qy        176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
              |||   :::||||||||||      |||   :::||||||:::::::::|||   |||
Db        664 ACAAACCTGAATGTGGAGCAAGTGTTTTTCTCCATCGCACGTGACATCAAGCAGAGGCTT 723

Qy        196 GlyHisGlnAlaAspGluPro---------SerPheGlnIleArgAspTyrValGluSer 212
                        :::|||         :::   :::|||                :::
Db        724 GCGGAGACTGATTCAAAGCCAGAGGACAAGGCAATCAAAATTAACAAGCCAGACCAGGGT 783

Qy        213 GlnLys------------LysArgSerSerCysCys 220
                 :::             :::||||||:::||||||
Db        784 TCCGAGGCACCAGCTGCCCAGCGATCTGCTTGCTGC 819
```

Search completed: June 23, 2003, 19:50:43
Job time : 231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 19:02:43 ; Search time 2487 Seconds
(without alignments)
2609.540 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09817199/runat_18062003_145102_795/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199_@CGN_1_1_2496_@runat_18062003_145102_795 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	1700	9	BC016615	BC016615 Homo sapi
2	1150	100.0	1977	9	AK098068	AK098068 Homo sapi
3	1145	99.6	1116	6	AX236082	AX236082 Sequence
4	1081	94.0	690	10	AF233582	AF233582 Mus muscu
5	977	85.0	576	6	AX236084	AX236084 Sequence
6	780.5	67.9	1513	9	BC007681	BC007681 Homo sapi
7	757	65.8	1320	9	AB027137	AB027137 Homo sapi
8	754	65.6	1098	10	RNU18771	U18771 Rattus norv
9	753	65.5	573	9	AF498952	AF498952 Homo sapi
10	731	63.6	1340	6	AR062279	AR062279 Sequence
11	710.5	61.8	3105	9	AK054846	AK054846 Homo sapi
12	649.5	56.5	2536	3	AY061826	AY061826 Drosophil
13	521.5	45.3	1337	10	BC019990	BC019990 Mus muscu
14	518.5	45.1	1074	8	AB024994	AB024994 Cicer ari
15	517.5	45.0	2145	3	DDISAS1A	M34456 D.discoideu
c 16	517.5	45.0	10034	2	AC115606	AC115606 Dictyoste
17	516	44.9	638	10	S53270	S53270 MEL-RAS-rel
18	516	44.9	765	5	DYGORA2	M38391 Discopyge o
19	516	44.9	911	8	AY084345	AY084345 Arabidops
20	516	44.9	2048	9	BC002977	BC002977 Homo sapi
21	516	44.9	2818	9	AK025165	AK025165 Homo sapi
22	513.5	44.7	624	9	AF498943	AF498943 Homo sapi
23	513.5	44.7	660	9	HSMRAB8	X56741 H.sapiens m
24	513.5	44.7	740	10	RNU53475	U53475 Rattus norv
25	513.5	44.7	760	4	CFRAB8	X56385 Canine rab8
26	513.5	44.7	782	8	PEAGTPBP09	D12548 Pea mRNA fo
27	513.5	44.7	1265	9	AB038995	AB038995 Homo sapi
28	512.5	44.6	1080	8	BVRAB1	Z49152 B.vulgaris
29	512	44.5	1980	9	S53268	S53268 Homo sapien
30	511.5	44.5	2497	6	AX285074	AX285074 Sequence
31	511.5	44.5	2497	6	AX285080	AX285080 Sequence
32	511.5	44.5	2497	6	AX285089	AX285089 Sequence
33	511.5	44.5	2497	6	AX285091	AX285091 Sequence
34	510	44.3	919	8	DCJ001367	AJ001367 Daucus ca
35	509.5	44.3	1082	8	LJRAB8D	Z73947 L.japonicus
36	509	44.3	955	8	S66160	S66160 ric1-ras-re
37	509	44.3	1163	8	LJRAB8A	Z73944 L.japonicus
38	508.5	44.2	1613	3	AB006189	AB006189 Drosophil
39	508.5	44.2	2349	3	AY060425	AY060425 Drosophil
40	507.5	44.1	979	8	LJRAB1C	Z73932 L.japonicus
41	507	44.1	681	8	AY056351	AY056351 Arabidops
42	507	44.1	821	8	AY052333	AY052333 Arabidops
43	507	44.1	986	8	AB079023	AB079023 Nicotiana
44	507	44.1	1025	8	AY086756	AY086756 Arabidops
45	507	44.1	1033	8	AB079021	AB079021 Nicotiana

ALIGNMENTS

RESULT 1

```
BC016615
LOCUS       BC016615               1700 bp    mRNA    linear   PRI 05-NOV-2001
DEFINITION  Homo sapiens, Similar to RAB37, member of RAS oncogene family,
            clone MGC:21391 IMAGE:4520191, mRNA, complete cds.
ACCESSION   BC016615
VERSION     BC016615.1  GI:16741620
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1700)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 28 Row: j Column: 15
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES              Location/Qualifiers
     source           1. .1700
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="MGC:21391 IMAGE:4520191"
                      /tissue_type="Prostate, adenocarcinoma."
                      /clone_lib="NIH_MGC_91"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              3. .674
                      /codon_start=1
                      /product="Similar to RAB37, member of RAS oncogene family"
                      /protein_id="AAH16615.1"
                      /db_xref="GI:16741621"
                      /translation="MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTC
                      FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRD
                      AQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGET
                      LAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSS
                      CCSFM"
BASE COUNT      437 a    444 c    460 g    359 t
ORIGIN

Alignment Scores:
Pred. No.:                1.77e-108       Length:          1700
Score:                    1150.00         Matches:         223
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-817-199B-2 (1-223) x BC016615 (1-1700)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          3 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 62

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         63 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 122

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        123 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 182

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        183 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 242

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        243 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 302

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        303 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 362

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        363 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 422

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        423 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 482

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        483 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 542

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        543 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 602

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        603 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 662

Qy        221 SerPheMet 223
              |||||||||
Db        663 TCCTTCATG 671

RESULT 2
AK098068
LOCUS       AK098068               1977 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
            to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION   AK098068
VERSION     AK098068.1  GI:21758001
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens trachea cDNA to mRNA, clone_lib:TRACH2
            clone:TRACH2000540.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
            Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
            Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1977)
  AUTHORS   Isogai,T. and Yamamoto,J.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .1977
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TRACH2000540"
                     /tissue_type="trachea"
                     /clone_lib="TRACH2"
                     /note="cloning vector: pME18SFL3"
     CDS             57. .728
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC05227.1"
                     /db_xref="GI:21758002"
                     /translation="MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTC
                     FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRD
                     AQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGET
                     LAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSS
                     CCSFM"
BASE COUNT      442 a    560 c    539 g    436 t
ORIGIN

Alignment Scores:
Pred. No.:              2.16e-108       Length:         1977
Score:                  1150.00         Matches:        223
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-817-199B-2 (1-223) x AK098068 (1-1977)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 116

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 176

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 236

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 296

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 356

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 416

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 476

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        477 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 536

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 596

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 656

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 716

Qy        221 SerPheMet 223
              |||||||||
Db        717 TCCTTCATG 725

RESULT 3
AX236082
LOCUS       AX236082               1116 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 13 from Patent WO0164887.
ACCESSION   AX236082
VERSION     AX236082.1  GI:15795889
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1116)
  AUTHORS   Meyers,R.A.
  TITLE     32705, 23224, 27423, 32700, 32712, novel human g-proteins
  JOURNAL   Patent: WO 0164887-A 13 07-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1116
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             124. .699
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC88255.1"
                     /db_xref="GI:15795890"
                     /translation="MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG
                     VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR
                     DVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELK
                     YRAGHQADEPSFQIRDYVESQKKRSSCCSFM"
BASE COUNT      260 a    313 c    322 g    221 t
ORIGIN

Alignment Scores:
Pred. No.:              3.35e-108       Length:         1116
Score:                  1145.00         Matches:        222
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            99.57%          Indels:         0
DB:                     6               Gaps:           0

US-09-817-199B-2 (1-223) x AX236082 (1-1116)

Qy          2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 90

Qy         22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 150

Qy         42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 210

Qy         62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 270

Qy         82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 330

Qy        102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 390

Qy        122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 450

Qy        142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 510

Qy        162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 570

Qy        182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 630

Qy        202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 690

Qy        222 PheMet 223
                ||||||
Db        691 TTCATG 696

RESULT 4
AF233582
LOCUS       AF233582                690 bp    mRNA    linear   ROD 02-MAY-2000
DEFINITION  Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
ACCESSION   AF233582
VERSION     AF233582.1  GI:7677421
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   Masuda,E.S., Luo,Y., Young,C., Shen,M., Rossi,A.B., Huang,B.C.,
            Yu,S., Bennett,M.K., Payan,D.G. and Scheller,R.H.
  TITLE     Rab37 is a novel mast cell specific GTPase localized to secretory
            granules
  JOURNAL   FEBS Lett. 470 (1), 61-64 (2000)
  MEDLINE   20189834
   PUBMED   10722846
REFERENCE   2  (bases 1 to 690)
  AUTHORS   Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
            Avenue, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .690
                     /gene="Rab37"
     CDS             1. .672
                     /gene="Rab37"
                     /codon_start=1
                     /product="GTPase Rab37"
                     /protein_id="AAF67162.1"
                     /db_xref="GI:7677422"
                     /translation="MTGTPGAATAGDGEAPERSPPFSPNYDLTGKVMLLGDSGVGKTC
                     FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGARVKLQIWDTAGQERFRSVTHAYYRD
                     AQALLLLYDITNQSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADVSSERVIRSEDGET
                     LAREYGVPFMETSAKTGMNVELAFLAIAKELKYRAGRQPDEPSFQIRDYVESQKKRSS
                     CCSFV"
BASE COUNT      160 a    181 c    201 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              6.63e-102       Length:         690
Score:                  1081.00         Matches:        209
Percent Similarity:     96.41%          Conservative:   6
Best Local Similarity:  93.72%          Mismatches:     8
Query Match:            94.00%          Indels:         0
DB:                     10              Gaps:           0

US-09-817-199B-2 (1-223) x AF233582 (1-690)

Qy          1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
              ||||||||||||||||||||||          ||||||||||||||||||||||||||||||
Db          1 ATGACTGGCACACCAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCG 60

Qy         21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
              |||   ||||||:::||||||||||||||||||||||||||||||||||:::||||||
Db         61 CCCTTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTC 120

Qy         41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 180

Qy         61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        181 GCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAG 240

Qy         81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTAC 300

Qy        101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        301 CGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACAAC 360

Qy        121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATTATGCTT 420

Qy        141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        421 CTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACA 480

Qy        161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
              ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db        481 CTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTG 540

Qy        181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||  |||  |||
Db        541 GAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTGAT 600

Qy        201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 660

Qy        221 SerPheMet 223
              ||||||:::
Db        661 TCCTTTGTG 669

RESULT 5
AX236084
LOCUS       AX236084                576 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 15 from Patent WO0164887.
ACCESSION   AX236084
VERSION     AX236084.1  GI:15795891
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 576)
  AUTHORS  Meyers,R.A.
  TITLE    32705, 23224, 27423, 32700, 32712, novel human g-proteins
  JOURNAL  Patent: WO 0164887-A 15 07-SEP-2001;
           Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      141 a    148 c    167 g    120 t
ORIGIN

Alignment Scores:
Pred. No.:              2.48e-91        Length:         576
Score:                  977.00          Matches:        191
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            84.96%          Indels:         0
DB:                     6               Gaps:           0

US-09-817-199B-2 (1-223) x AX236084 (1-576)

Qy         33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60

Qy         53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy         73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy         93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        213 GlnLysLysArgSerSerCysCysSerPheMet 223
              |||||||||||||||||||||||||||||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATG 573

RESULT 6
BC007681
LOCUS       BC007681                1513 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, RAB26, member RAS oncogene family, clone MGC:3503
            IMAGE:3627067, mRNA, complete cds.
ACCESSION   BC007681
VERSION     BC007681.1  GI:14043378
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1513)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 12 Row: m Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 5931611.
FEATURES             Location/Qualifiers
     source          1. .1513
                     /organism="Homo sapiens"
                     /db_xref="LocusID:25837"
                     /db_xref="taxon:9606"
                     /clone="MGC:3503 IMAGE:3627067"
                     /tissue_type="Uterus, endometrium adenocarcinoma"
                     /clone_lib="NIH_MGC_44"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             187. .759
                     /codon_start=1
                     /product="RAB26, member RAS oncogene family"
                     /protein_id="AAH07681.1"
                     /db_xref="GI:14043379"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      326 a    452 c    449 g    286 t
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-70        Length:         1513
Score:                  780.50          Matches:        153
Percent Similarity:     79.57%          Conservative:   30
Best Local Similarity:  66.52%          Mismatches:     32
Query Match:            67.87%          Indels:         16
DB:                     9               Gaps:           3

US-09-817-199B-2 (1-223) x BC007681 (1-1513)

Qy          6 GlyAlaValAlaThrArgAspGlyGluAla---------ProGluArgSerPro------ 20
              ||||||          |||   |||   |||         |||:::      |||
Db         62 GGGGCCCGACCGGCGCGCTCCGGGACTGCGCTTTCCGGCCCCGACGCGCCGCCCAACGGG 121

Qy         21 ProCysSerProSer----------------------------TyrAspLeuThrGly 30
              ||||||||||||                              ||||||:::
Db        122 CCTTGCAGCCCGGC-CGGCCCTCGCTTGGCGGCGGTGTCGACTTCTACGACGTCGCCTTC 180
```

```
Qy         31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
              ||||||||||||:::||||||:::||||||||||||||||||   |||:::::::||||||
Db        181 AAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTGCTGGTGCGATTCAAG 240

Qy         51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
              ||||||||||||||:::||||||||||||:::||||||||||||||||||||||||||||
Db        241 GATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAA 300

Qy         71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
              |||:::   ||||||||||||:::||||||||||||:::||||||||||||||||||||
Db        301 GTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGGAG 360

Qy         91 ArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyr 110
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        361 CGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCTCTGCTGCTGCTCTAC 420

Qy        111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
              |||:::||||||||||:::||||||||||||||:::||||||||||||||||||||||||
Db        421 GATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTGACCGAGATCCACGAG 480

Qy        131 TyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
              |||||||||   ||||||   :::||||||||||||||||||   |||   :::   |||
Db        481 TACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTGGACTCTGCCCATGAG 540

Qy        151 ArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeu 170
              ||||||::::::   ||||||||||||   ||||||:::||||||||:::||||||:::
Db        541 CGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTATGGACTGCCCTTCATG 600

Qy        171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
              ||||||||||||||||||||||:::||||||:::||||||||   ||||||||||||||||
Db        601 GAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCACAGCCATAGCAAAGGAG 660

Qy        191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
              ||||||   |||:::                ||||||   |||::::::   |||||||||
Db        661 TTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCGGCTGCATGATTACGTT 720

Qy        211 GluSerGlnLysLysArgSerSerCysCys 220
              :::  :::   :::   :::||||||||||
Db        721 AAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 750

RESULT 7
AB027137
LOCUS       AB027137               1320 bp    mRNA    linear   PRI 29-SEP-1999
DEFINITION  Homo sapiens v46133 mRNA for RAB-26, complete cds.
ACCESSION   AB027137
VERSION     AB027137.1  GI:5931611
KEYWORDS    RAB-26; v46133.
SOURCE      Homo sapiens cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S.,
            Muramatsu,M. and Saito,T.
  TITLE     Human RAS-related protein RAB-26
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S.,
            Muramatsu,M. and Saito,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-1999) Toshiyuki Saito, National Institute of
            Radiological Sciences, Genome Research Group; Inage-ku Anagawa
            4-9-1, Chiba, Chiba 263-8555, Japan (E-mail:t_saito@nirs.go.jp,
            Tel:81-43-201-3135, Fax:81-43-251-9818)
FEATURES             Location/Qualifiers
     source          1. .1320
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1320
                     /gene="v46133"
     CDS             4. .576
                     /gene="v46133"
                     /note="RAS-related protein"
                     /codon_start=1
                     /product="RAB-26"
                     /protein_id="BAA84707.1"
                     /db_xref="GI:5931612"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      273 a    379 c    395 g    273 t
ORIGIN

Alignment Scores:
Pred. No.:              2.78e-68        Length:         1320
Score:                  757.00          Matches:        139
Percent Similarity:     88.36%          Conservative:   28
Best Local Similarity:  73.54%          Mismatches:     22
Query Match:            65.83%          Indels:         0
DB:                     9               Gaps:           0

US-09-817-199B-2 (1-223) x AB027137 (1-1320)

Qy         32 ValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAsp 51
              |||||||||:::||||||:::||||||||||||||||||   |||:::::::|||||||||
Db          1 GTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTGCTGGTGCGATTCAAGGAT 60

Qy         52 GlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysVal 71
              |||||||||||||:::||||||||||||:::||||||||||||||||||||||||||||||
Db         61 GGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAAGTT 120

Qy         72 ValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArg 91
              :::   ||||||||||||:::||||||||||||:::||||||||||||||||||||||||
Db        121 CTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGGAGCGG 180

Qy         92 PheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAsp 111
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        181 TTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCTCTGCTGCTGCTCTACGAT 240

Qy        112 IleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyr 131
              :::||||||||||:::||||||||||||||||:::||||||||||||||||||||||||||
Db        241 GTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTGACCGAGATCCACGAGTAC 300

Qy        132 AlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArg 151
              ||||||   ||||||   :::||||||||||||||||||   |||   :::   ||||||
Db        301 GCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGT 360

Qy        152 ValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGlu 171
              |||:::::::   ||||||||||||   ||||||:::||||||||:::||||||:::|||
Db        361 GTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTATGGACTGCCCTTCATGGAG 420

Qy        172 ThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeu 191
              ||||||||||||||||||:::||||||:::||||||||   ||||||||||||||||||||
Db        421 ACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCACAGCCATAGCAAAGGAGTTG 480

Qy        192 LysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGlu 211
              |||   |||:::                ||||||   |||::::::   ||||||||||:::
Db        481 AAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCGGCTGCATGATTACGTTAAG 540

Qy        212 SerGlnLysLysArgSerSerCysCys 220
                 :::   :::   :::||||||||||
Db        541 AGGGAGGGTCGAGGGGCCTCCTGCTGC 567

RESULT 8
RNU18771
LOCUS       RNU18771               1098 bp    mRNA    linear   ROD 18-JUL-1995
DEFINITION  Rattus norvegicus Rab26 mRNA, complete cds.
ACCESSION   U18771
VERSION     U18771.1  GI:619733
KEYWORDS    .
SOURCE      rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
```

```
          Rattus.
REFERENCE 1  (bases 1 to 1098)
  AUTHORS Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
  TITLE   Molecular cloning of a new member of the Rab protein family, Rab
          26, from rat pancreas
  JOURNAL Biochem. Biophys. Res. Commun. 207 (3), 950-956 (1995)
  MEDLINE 95169156
   PUBMED 7864900
REFERENCE 2  (bases 1 to 1098)
  AUTHORS Williams,J.A.
  TITLE   Direct Submission
  JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of
          Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .1098
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue_type="pancreas"
                     /dev_stage="adult"
     5'UTR           1. .29
     CDS             30. .602
                     /codon_start=1
                     /product="Rab26"
                     /protein_id="AAA69955.1"
                     /db_xref="GI:619734"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     MKVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQ
                     DVVLMLLGNKVDSTQERVVKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELK
                     QRSTKAPSEPRFRLHDYVKREGRGVSCCRL"
     3'UTR           603. .1098
     polyA_signal    1063. .1068
BASE COUNT      267 a    277 c    306 g    248 t
ORIGIN

Alignment Scores:
Pred. No.:                4.44e-68       Length:          1098
Score:                    754.00         Matches:         140
Percent Similarity:       87.89%         Conservative:    27
Best Local Similarity:    73.68%         Mismatches:      23
Query Match:              65.57%         Indels:          0
DB:                       10             Gaps:            0

US-09-817-199B-2 (1-223) x RNU18771 (1-1098)

Qy         31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
              ||||||||||||:::|||||:::||||||||||||||||||   |||::::::||||||
Db         24 AAGGTCATGCTGGTGGGGGATTCCGGTGTGGGGAAGACGTGCCTGCTCGTGCGCTTCAAG 83

Qy         51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
              ||||||||||||||:::||||||||||:::||||||||||||||||||||||||||||||
Db         84 GATGGGGCCTTCCTGGCTGGTACCTTCATCTCCACTGTGGGCATCGACTTCCGGAATAAA 143

Qy         71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
              |||:::   |||||||||:::::::||||||||||||||||||||||||||||||||||
Db        144 GTTCTGGATGTGGATGGCATGAAGGTGAAGCTGCAGATCTGGGACACGGCTGGTCAGGAG 203

Qy         91 ArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyr 110
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        204 CGGTTCCGAAGTGTCACCCATGCCTACTACCGGGATGCTCACGCACTGTTGCTGCTCTAC 263

Qy        111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
              ||||||||||||||   ||||||||||||||:::||||||||||||||||||   |||
Db        264 GACATCACCAACAAAGATTCCTTCGACAACATCCAGGCCTGGTTGACAGAAATCCAGGAA 323

Qy        131 TyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
              |||||||||:::||||||||:::||||||||||||||||||   |||   :::   |||
Db        324 TATGCCCAGCAGGACGTGGTGCTCATGCTGTTGGGGAACAAGGTTGACTCTACTCAAGAA 383

Qy        151 ArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeu 170
              ||||||::::::   ||||||||||||   ||||||:::||||||||:::|||||||:::
Db        384 CGAGTGGTAAAGAGGGAAGATGGGGAGAAATTAGCCAAGGAGTATGGGCTGCCATTCATG 443

Qy        171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
              ||||||||||||||||:::|||:::||||||:::||||||||||   ||||||||||||||
Db        444 GAGACCAGCGCCAAGAGCGGCCTCAATGTGGACTTGGCTTTTACAGCCATAGCAAAGGAG 503

Qy        191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
              ||||||   |||:::              ||||||   |||::::::   |||||||||
Db        504 CTGAAACAAAGATCCACCAAGGCTCCCAGTGAGCCGCGCTTCAGGCTGCATGACTATGTG 563

Qy        211 GluSerGlnLysLysArgSerSerCysCys 220
              :::   :::   :::      |||||||||
Db        564 AAGAGGGAGGGCCGAGGGGTCTCCTGCTGT 593

RESULT 9
AF498952
LOCUS       AF498952                573 bp    mRNA    linear   PRI 01-MAY-2002
DEFINITION  Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete
            cds.
ACCESSION   AF498952
VERSION     AF498952.1  GI:20379079
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
  TITLE     Homo sapiens RAB family small GTP binding protein RAB26
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 573)
  AUTHORS   Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
            Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     gene            1. .573
                     /gene="RAB26"
     CDS             1. .573
                     /gene="RAB26"
                     /note="RAB family member"
                     /codon_start=1
                     /product="small GTP binding protein RAB26"
                     /protein_id="AAM21100.1"
                     /db_xref="GI:20379080"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      117 a    151 c    191 g    114 t
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-68       Length:          573
Score:                    753.00         Matches:         138
Percent Similarity:       88.30%         Conservative:    28
Best Local Similarity:    73.40%         Mismatches:      22
Query Match:              65.48%         Indels:          0
DB:                       9              Gaps:            0

US-09-817-199B-2 (1-223) x AF498952 (1-573)

Qy         33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
              |||||:::|||||:::||||||||||||||||||   |||:::::|||||||||||
Db          1 ATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTGCTGGTGGCGATTCAAGGATGGT 60

Qy         53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
              |||||||||:::||||||||||||:::|||||||||||||||||||||||||||||:::
Db         61 GCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAAGTTCTG 120
```

```
Qy         73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
              ||||||||||||:::|||||||||||||:::|||||||||||||||||||||||||||||
Db        121 GACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGGAGCGGTTC 180

Qy         93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIle 112
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||:::
Db        181 CGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCTCTGCTGCTGCTCTACGATGTC 240

Qy        113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
              |||||||:::|||||||||||||:::|||||||||||||||||||||||||||||||||
Db        241 ACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTGACCGAGATCCACGAGTACGCC 300

Qy        133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
              |||   ||||||   :::|||||||||||||||||||   |||   :::   |||||||||
Db        301 CAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGTGTG 360

Qy        153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
              ::::::   ||||||||||||   |||||||:::|||||||||:::|||||||:::||||||
Db        361 GTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTATGGACTGCCCTTCATGGAGACC 420

Qy        173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
              ||||||||||||||||:::|||||||:::|||||||||   ||||||||||||||||||||
Db        421 AGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCACAGCCATAGCAAAGGAGTTGAAG 480

Qy        193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
              |||:::                |||||   |||::::::   |||||||||:::
Db        481 CAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCGGCTGCATGATTACGTTAAGAGG 540

Qy        213 GlnLysLysArgSerSerCysCys 220
              :::   :::   :::||||||||||
Db        541 GAGGGTCGAGGGGCCTCCTGCTGC 564

RESULT 10
AR062279
LOCUS       AR062279               1340 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5843717.
ACCESSION   AR062279
VERSION     AR062279.1  GI:5989970
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1340)
  AUTHORS   Hillman,J.L. and Guegler,K.J.
  TITLE     Rab protein
  JOURNAL   Patent: US 5843717-A 2 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .1340
                     /organism="unknown"
BASE COUNT      276 a    387 c    403 g    274 t
ORIGIN

Alignment Scores:
Pred. No.:              1.32e-65        Length:         1340
Score:                  731.00          Matches:        141
Percent Similarity:     86.73%          Conservative:   29
Best Local Similarity:  71.94%          Mismatches:     25
Query Match:            63.57%          Indels:         2
DB:                     6               Gaps:           0

US-09-817-199B-2 (1-223) x AR062279 (1-1340)

Qy         26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
              ||||||:::      ||||||||||||:::|||||||:::|||||||||||||||||||
Db         22 TACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTG 81

Qy         46 Leu-IleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIl 65
              ||| ::::::||||||||||||||||||||||:::|||||||||||:::||||||   ||
Db         82 CTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGC-AT 140

Qy         65 eAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAs 85
              ||||||||||||||||||:::   ||||||||||||:::|||||||||||||:::|||||
Db        141 TGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGA 200

Qy         85 pThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAl 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        201 CACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGC 260

Qy        105 aLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLe 125
              ||||||||||||||||||:::|||||||||:::|||||||||||||||||:::||||||||
Db        261 TCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCT 320

Qy        125 uThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAl 145
              |||||||||||||||||||||||   ||||||   :::|||||||||||||||||||||
Db        321 GACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGT 380

Qy        145 aAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTy 165
              |||   :::   ||||||||||::::::   ||||||||||||   |||||||:::|||||
Db        381 GGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTA 440

Qy        165 rGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLe 185
              ||||:::|||||||:::|||||||||||||||||||||:::|||||||:::||||||||||
Db        441 TGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCAC 500

Qy        185 uAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGl 205
              ||||||||||||||||||||   |||:::                ||||||   |||::
Db        501 AGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCG 560

Qy        205 nIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
              ::::   ||||||||||:::   :::   :::   :::||||||||||
Db        561 GCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606

RESULT 11
AK054846
LOCUS       AK054846               3105 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately
            similar to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION   AK054846
VERSION     AK054846.1  GI:16549463
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
            clone:BRACE2002812.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3105)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .3105
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BRACE2002812"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE2"
```

```
                     /note="cloning vector: pME18SFL3"
BASE COUNT      680 a    905 c    898 g    622 t
ORIGIN

Alignment Scores:
Pred. No.:                5e-63          Length:         3105
Score:                    710.50         Matches:        182
Percent Similarity:       39.47%         Conservative:   13
Best Local Similarity:    36.84%         Mismatches:     26
Query Match:              61.78%         Indels:         274
DB:                       9              Gaps:           5

US-09-817-199B-2 (1-223) x AK054846 (1-3105)

Qy          3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro--- 21
              |||      ||||||   :::::::::   |||         :::|||
Db        375 GGGAGAGCCGGAGCGGCGAGCTCCAAGCCTGGCATGGACCTGCAGAGACCCGATTCCTAC 434

Qy         22 ---------CysSerProSerTyrAsp------LeuThrGlyLysValMetLeuLeuGly 36
                       |||   ::::::      :::      |||   :::|||:::|||
Db        435 CAGGGAGGAGCTGGCCCTGACTTCAACGACCACGTCCTGCATAAGACCATCCTGGTGGGT 494

Qy         37 AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer 56
              |||:::|||||||||||||||      |||:::||||||      |||   |||:::
Db        495 GACAGTGGTGTGGGAAAGACGTCTCTGCTGGTTCAGTTCGATCAGGGCAAGTTCATCCCC 554

Qy         57 GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGly 76
              |||:::|||   |||||||||||||||   |||   ||||||||||||||||||||||
Db        555 GGCTCCTTCTCGGCCACTGTGGGCATCGGATTCACGAACAAGGTGGTGACTGTGGATGGC 614

Qy         77 ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 GTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACC 674

Qy         97 HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 CATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCT 734

Qy        117 SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 TCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTG 794

Qy        137 ValIleMetLeuLeuGlyAsnLys---------------------------------- 144
              |||||||||||||||||||||||
Db        795 GTGATCATGCTGCTAGGCAACAAGGTGAGTGGCTCCGGGGCAGGGTCAGCCCAGCCCTGC 854

Qy        144 ---------------------------------------------------------- 144

Db        855 ACTTCCTCAGCCCTAGCCGGCCCCATAACCACCCAAGAACAGTTATCTAGGCATCCTTCC 914

Qy        144 ---------------------------------------------------------- 144

Db        915 TGAAAAGGACTCTGCAGCCTCCAGCTCAGGGGTCAGACATATCTGGAGGCTTCTGCCCAT 974

Qy        144 ---------------------------------------------------------- 144

Db        975 CCCATCTGCCCCTTCCAGGGAAAGTCCAAGTTGTTGCCTGAGAAATCAAGGGGTGCCCAG 1034

Qy        144 ---------------------------------------------------------- 144

Db       1035 TTCTCAGCCCCCATTAGAGCAGAGTGAACAGGGTCCCAGGTCAGGGGCTAAGAGTGCAAA 1094

Qy        144 ---------------------------------------------------------- 144

Db       1095 GGGTTAGCCCCAACTGCTGTCCTATTCCAAGACCCTTTACCAAAGGTGAGATCCCAGAGC 1154

Qy        144 ---------------------------------------------------------- 144

Db       1155 TGGGAGCTACACTGGGCAGAAACCCTGGCCCCAGGCCAATCACACCTGCCTGCAGTCCCT 1214

Qy        144 ---------------------------------------------------------- 144

Db       1215 TGGGCCACCAGCAGAGGGCAGGCAACGCCTGCTTCTGGGGCAAAATATGGGCCCGCTGGG 1274

Qy        145 ----------------------------------------AlaAspMetSerSerGl 150
                                                     ||||||||||||||||
Db       1275 GCGGAGGCCTCCTTCCCCAGAGTGACCCATTTGGGCTTGACAGGCGGATATGAGCAGCGA 1334

Qy        150 uArgValIleArgSerGluAspGlyGluThrLeuAlaArg------------------- 163
               ||||||||||||||||||||||||||||||||||||||||
Db       1335 AAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGTAAGTGATTGTCTGTGGGA 1394

Qy        164 ---------------------------------------------------------Gl 164
                                                                        ||
Db       1395 CAGGGTGAAGGGTGGGGGCAACCCGACGCTGGCCCTGAGGACACTCTCTCCCGGGCAGGA 1454

Qy        164 uTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPh 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTT 1514

Qy        184 eLeuAlaIleAlaLys---------------------------------------- 189
               |||||||||||||
Db       1515 TCTGGCCATCGCCAA-GTGAGAGCTGGGCAGGGAAGGGAAGTGTGCGGGGCAGGGCGGCA 1573

Qy        189 ---------------------------------------------------------- 189

Db       1574 CACTCCAGGAATCCAGTAGGGCCCGGCCCCTGGCCCAGCCCCTGGACACACCTGCATTCT 1633

Qy        189 ---------------------------------------------------------- 189

Db       1634 GCAGGCTGAGGTCCATTTGCTCTGGGAGCACTGGGCCACTGGGAGAGGGGAGGGGCGGC 1693

Qy        190 ---------------------------------------------------------Gl 190
                                                                        ||
Db       1694 TCAGCTCCTCACCCCAGCCCAGCCCAGCCCAGCCCAGCCCATTGTCTCTTCTTCAAGGGA 1753

Qy        190 uLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVa 210
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 ACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGT 1813

Qy        210 lGluSerGlnLysLysArgSerSerCysCysSerPheMet 223
               ||||||||||||||||||||||||||||||||||||||
Db       1814 AGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATG 1853

RESULT 12
AY061826
LOCUS       AY061826               2536 bp    mRNA    linear   INV 10-NOV-2001
DEFINITION  Drosophila melanogaster GH21984 full length cDNA.
ACCESSION   AY061826
VERSION     AY061826.1  GI:16902019
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 2536)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
```

```
          artifacts associated with the generation of cDNA clones that may
          have not been detected in our initial analyses such as internal
          priming, priming from contaminating genomic DNA, retained introns
          due to reverse transcription of unspliced precursor RNAs, and
          reverse transcriptase errors that result in single base changes.
          For further information about this sequence, including its location
          and relationship to other sequences, please visit our Web site
          (http://fruitfly.berkeley.edu) or send email to
          cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .2536
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="78C9-78D1"
     gene            1. .2536
                     /gene="CG7605"
                     /note="alignment with genomic scaffold AE003594"
                     /db_xref="FLYBASE:FBgn0037072"
     CDS             1172. .2338
                     /gene="CG7605"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="GH21984p"
                     /protein_id="AAL27637.1"
                     /db_xref="GI:16902020"
                     /db_xref="FLYBASE:FBgn0037072"
                     /translation="MASTAVGLGGGEGDPGAGGPPAGSAHPDDASSMSDDVFEDAETT
                     QARIEELRRRPFGDGSYNPPAAPASVSASITTTTTQQQQQHHNPSHHHQSSHHQPSHH
                     HHHHHHSQLSLTGSHHYHDDAIMAPVQRSATGYPGYRPSREAMQMYAYGTDDYDDDYN
                     DGWRSYRYDEVDMHPAPSNAHQQPFDDTVNHKTILLGDSGVGKTSFLVKYNTGEFRLG
                     SFSATVGIALTNKVVVVDGTRVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDVTNK
                     TTYDNIRAWLGEIREYAQEDVVIVLIGNKADCSGSERQVKREDGERLGREHNVPFMET
                     SAKTGLNVELSFTAVARQLKSRGYEHGDDGRFNVHDFVRDNTKARSVCAQCRNM"
BASE COUNT      831 a    621 c    624 g    460 t
ORIGIN

Alignment Scores:
Pred. No.:                7.01e-57       Length:          2536
Score:                    649.50         Matches:         132
Percent Similarity:       75.96%         Conservative:    26
Best Local Similarity:    63.46%         Mismatches:      48
Query Match:              56.48%         Indels:          3
DB:                       3              Gaps:            2

US-09-817-199B-2 (1-223) x AY061826 (1-2536)

Qy         16 ProGluArgSerProProCysSerProSer---TyrAspLeuThrGlyLysValMetLeu 34
              |||      :::|||      |||||||||       ||||||   |||   :::|||
Db       1698 CCCCGTCCAACGCCCATCAGCAGCCCTTCGATGATACGGTTAACCAC-AAGACGATCCTG 1756

Qy         35 LeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPhe 54
              |||||||||:::||||||||||||||||   ||||||::::::::      |||   |||
Db       1757 CTGGGGGATTCGGGCGTGGGAAAGACCTCCTTCCTGGTCAAATACAATACGGGCGAGTTC 1816

Qy         55 LeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrVal 74
                    |||:::|||   |||||||||||||||         ||||||||||   |||
Db       1817 CGACTGGGTTCCTTTTCCGCCACAGTGGGCATTGCGCTTACGAACAAAGTGGTGGTCGTC 1876

Qy         75 AspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSer 94
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 GATGGAACGCGCGTCAAGCTGCAAATCTGGGACACAGCTGGTCAGGAGCGATTCCGGAGC 1936

Qy         95 ValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsn 114
              ||||||||||||||||||||||||||   |||||||||||||||||||:::||||||
Db       1937 GTTACCCACGCCTATTATCGCGACGCGCACGCTCTACTGCTGCTGTACGACGTGACCAAC 1996

Qy        115 LysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArg 134
              |||:::::::::|||||||||||||||||||   ||||||   ||||||||||||
Db       1997 AAGACCACCTATGACAACATTCGCGCCTGGCTGGGCGAGATCCGGGAGTACGCGCAGGAG 2056

Qy        135 AspValValIleMetLeuLeuGlyAsnLysAlaAspMetSer---SerGluArgValIle 153
              ||||||||||||:::|||:::|||||||||||||||||   |||   |||||||||   :::

Db       2057 GACGTGGTCATCGTTTTAATAGGCAACAAGGCCGACTGCAGCGGCAGCGAGCGGCAGGTG 2116

Qy         154 ArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSer 173
               :::   ||||||||||||   |||   ||||||:::   |||||||||:::||||||||
Db       2117 AAGCGGGAGGATGGGGAGCGTTTGGGGCGGGAGCACAACGTGCCCTTCATGGAGACCTCG 2176

Qy         174 AlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyr 193
               ||||||||||||:::||||||||||||:::|||   |||:::|||:::::::||||||
Db       2177 GCCAAGACGGGACTCAATGTGGAGCTGTCCTTCACAGCGGTGGCCAGGCAACTAAAGAGT 2236

Qy         194 ArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGln 213
               |||               |||:::      |||   :::   |||:::|||
Db       2237 CGCGGCTACGAGCACGGCGATGATGGAAGGTTCAATGTGCATGATTTTGTGCGTGACAAT 2296

Qy         214 LysLysArgSerSerCysCysSer 221
                  |||      |||   |||:::
Db       2297 ACAAAGGCGCGCTCGGTTTGCGCC 2320

RESULT 13
BC019990
LOCUS       BC019990                1337 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to mel transforming oncogene (derived from
            cell line NK14)- RAB8 homolog, clone MGC:28462 IMAGE:4161110, mRNA,
            complete cds.
ACCESSION   BC019990
VERSION     BC019990.1  GI:18043408
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1337)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 37 Row: c Column: 17
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES             Location/Qualifiers
     source          1. .1337
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:28462 IMAGE:4161110"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             1. .624
                     /codon_start=1
                     /product="Similar to mel transforming oncogene (derived
                     from cell line NK14)- RAB8 homolog"
```

```
                     /protein_id="AAH19990.1"
                     /db_xref="GI:18043409"
                     /translation="MAKTYDYLFKLLLIGDSGVGKTCVLFRFSEDAFNSTFISTIGID
                     FKIRTIELDGKRIKLQIWDTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNW
                     IRNIEEHASADVEKMILGNKCDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVEN
                     AFFTLARDIKAKMDKKLEGNSPQGSSHGVKITVEQQKRTSFFRCSLL"
BASE COUNT      367 a    338 c    341 g    291 t
ORIGIN

Alignment Scores:
Pred. No.:              4.19e-44        Length:         1337
Score:                  521.50          Matches:        101
Percent Similarity:     69.76%          Conservative:   42
Best Local Similarity:  49.27%          Mismatches:     55
Query Match:            45.35%          Indels:         7
DB:                     10              Gaps:           3

US-09-817-199B-2 (1-223) x BC019990 (1-1337)

Qy         25 SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCys 44
              :::||||||       |||:::::::|||:::||||||::|||||||||||||||||||
Db         10 ACCTACGATTACCTGTTCAAGCTGCTGCTGATCGGGGACTCGGGGGTAGGGAAGACCTGT 69

Qy         45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64
                 |||   :::|||   :::   ||||||   :::   ||||||||||:::|||:::|||
Db         70 GTCCTGTTCCGCTTCTCCGAGGACGCCTTC---AACTCCACATTCATCTCTACCATAGGA 126

Qy         65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrp 84
              |||||||||:::   :::   :::   :::||||||   |||:::|||||||||||||||
Db        127 ATTGACTTTAAAATTAGGACCATAGAGCTCGATGGCAAGAGGATTAAACTGCAGATATGG 186

Qy         85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGln 104
              ||||||||||||||||||||||||||:::::|||  ||||||||||||   |||
Db        187 GACACGGCCGGCCAGGAGCGGTTTCGAACAATCACGACAGCCTACTACAGGGGTGCCATG 246

Qy        105 AlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrp 124
              ::::::|||:::||||||||||||||::   ||||||||||||||||   |||
Db        247 GGTATCATGCTGGTCTACGACATTACCAATGAGAAGTCCTTTGACAACATCCGGAATTGG 306

Qy        125 LeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLys 144
              :::      |||   |||:::|||      ||||||      |||:::||||||||||||
Db        307 ATTCGGAACATTGAAGAGCATGCCTCTGCAGACGTGGAGAAGATGATACTGGGGAATAAG 366

Qy        145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
              |||:::::::   :::|||   :::      |||   ||||||   ||||||   :::
Db        367 TGTGATGTGAATGACAAGAGACAGGTGTCCAAGGAACGGGGAGAAAAGCTGGCACTCGAC 426

Qy        165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
              ||||||:::   |||:::|||||||||||||||||      :::|||||||||   ||||||
Db        427 TATGGGATCAAGTTCATGGAGACCAGTGCAAAGGCCAACATTAATGTGGAGAATGCATTT 486

Qy        185 LeuAlaIleAlaLysGluLeuLysTyrArg--------------AlaGlyHisGlnAla 199
                    :::|||:::::::::|||   :::                  |||:::
Db        487 TTCACTCTTGCCAGGGATATCAAAGCAAAAATGGACAAAAAATTGGAAGGGAACAGCCCG 546

Qy        200 AspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCys 219
                    |||      ::::::      ||||||   ||||||:::   |||
Db        547 CAGGGGAGCAGCCATGGAGTCAAGATCACAGTGGAGCAGCAGAAGAGGACCAGCTTCTTC 606

Qy        220 ---CysSerPheMet 223
                 ||||||   :::
Db        607 CGGTGCAGTCTCCTG 621

RESULT 14
AB024994
LOCUS       AB024994               1074 bp    mRNA    linear   PLN 02-APR-1999
DEFINITION  Cicer arietinum mRNA for rab-type small GTP-binding protein,
            complete cds.
ACCESSION   AB024994
VERSION     AB024994.1  GI:4586579
KEYWORDS    rab-type small GTP-binding protein.
SOURCE      Cicer arietinum (strain:ILC3279) leaf cDNA to mRNA, clone:INR134.
  ORGANISM  Cicer arietinum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
            Cicer.
REFERENCE   1  (sites)
  AUTHORS   Ichinose,Y., Tiemann,K., Schwenger-Erger,C., Toyoda,K., Hein,F.,
            Hanselle,T. and Barz,W.
  TITLE     Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and
            Elicited Cell Cultures as Detected by Differential
            cDNA-Hybridization
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1074)
  AUTHORS   Ichinose,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1999) Yuki Ichinose, Okayama University, Faculty
            of Agriculture; Tsushima-naka, 1-1-1, Okayama, Okayama 700-8530,
            Japan (E-mail:yuki@cc.okayama-u.ac.jp, Tel:+81-86-251-8308,
            Fax:+81-86-251-8308)
FEATURES             Location/Qualifiers
     source          1. .1074
                     /organism="Cicer arietinum"
                     /strain="ILC3279"
                     /db_xref="taxon:3827"
                     /clone="INR134"
                     /tissue_type="leaf"
     CDS             107. .715
                     /codon_start=1
                     /product="rab-type small GTP-binding protein"
                     /protein_id="BAA76422.1"
                     /db_xref="GI:4586580"
                     /translation="MNPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLDSYISTIGVD
                     FKIRTVEQDGKTIKLQIWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQW
                     LNEIDRYASENVNKLLVGNKCDLAANKVVSSETAKAFADEIGIPFMETSAKNATNVEQ
                     AFMAMAAEIKNRMASQPANNARPPTVQIRGQPMNQKSGCCST"
     polyA_site      1074
                     /note="8 A nucleotides"
BASE COUNT      310 a    223 c    207 g    334 t
ORIGIN

Alignment Scores:
Pred. No.:              6.4e-44         Length:         1074
Score:                  518.50          Matches:        97
Percent Similarity:     69.90%          Conservative:   47
Best Local Similarity:  47.09%          Mismatches:     55
Query Match:            45.09%          Indels:         7
DB:                     8               Gaps:           3

US-09-817-199B-2 (1-223) x AB024994 (1-1074)

Qy         20 ProProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39
              |||      ::::|||   ||||||        |||:::::::|||:::|||||||:::|||
Db        101 CCCGTCATGAATCCCGAATATGACTATTTGTTCAAGCTTTTGTTGATTGGAGATTCTGGT 160

Qy         40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59
              |||||||||:::|||   |||::::::|||   |||   ::::::|||      ::::::
Db        161 GTGGGCAAGTCATGTCTCCTCCTGAGGTTTGCTGATGATTCATACCTTGAC---AGCTAT 217

Qy         60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79
              |||:::|||:::|||:::|||||||:::   :::   |||      ||||||      :::
Db        218 ATCAGTACAATTGGAGTGGACTTTAAAATTCGCACTGTTGAGCAAGACGGGAAGACCATT 277

Qy         80 LysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99
              ||||||||||||||||||||||||||||||||||||||||:::::|||   :::|||
Db        278 AAACTTCAAATTTGGGACACTGCTGGTCAAGAACGTTTCCGGACTATCACTAGCAGCTAC 337

Qy        100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAsp 119
              ||||||   |||      ::::::::::|||||||:::|||:::::   ||||||:::
Db        338 TATCGTGGGGCTCATGGCATAATTGTTGTTTATGATGTCACTGACCAAGAGAGCTTTAAC 397

Qy        120 AsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMet 139
              |||::::::   ||||||   ||||||      ||||||      :::|||      :::
Db        398 AATGTTAAGCAGTGGCTGAATGAAATTGACCGTTATGCAAGTGAAAATGTAAACAAGCTT 457
```

```
Qy        140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159
              |||:::|||||||||   |||:::::::::   :::|||:::   ||||||      :::
Db        458 CTAGTTGGAAACAAGTGTGATCTCGCGGCAAATAAAGTTGTGTCCTCTGAAACAGCAAAG 517

Qy        160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
                  |||   |||   |||:::||||||:::|||||||||||||||      |||
Db        518 GCATTTGCAGATGAAATTGGAATTCCTTTCATGGAAACTAGTGCTAAAAATGCCACCAAC 577

Qy        180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGln--- 198
              ||||||   ||||||:::|||:::|||   |||:::|||   |||         |||
Db        578 GTGGAACAGGCTTTCATGGCCATGGCTGCTGAAATTAAAAACAGAATGGCAAGTCAACCA 637

Qy        199 ---------AlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLys 215
                      |||      |||:::   |||||||||            :::||||||
Db        638 GCGAACAATGCCAGGCCTCCAACAGTCCAAATTCGAGGACAGCCAATGAACCAAAAG--- 694

Qy        216 ArgSerSerCysCysSer 221
                |||   |||||||||
Db        695 ---TCTGGTTGCTGCTCG 709

RESULT 15
DDISAS1A
LOCUS       DDISAS1A                2145 bp    mRNA    linear   INV 27-APR-1993
DEFINITION  D.discoideum GTP-binding protein (SAS1) gene, complete cds.
ACCESSION   M34456
VERSION     M34456.1  GI:167874
KEYWORDS    GTP-binding protein.
SOURCE      D.discoideum, cDNA to mRNA.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 2145)
  AUTHORS   Saxe,S.A. and Kimmel,A.R.
  TITLE     SAS1 and SAS2, GTP-binding protein genes in Dictyostelium
            discoideum with sequence similarities to essential genes in
            Saccharomyces cerevisiae
  JOURNAL   Mol. Cell. Biol. 10 (5), 2367-2378 (1990)
  MEDLINE   90220623
   PUBMED   2109188
FEATURES             Location/Qualifiers
     source          1. .2145
                     /organism="Dictyostelium discoideum"
                     /db_xref="taxon:44689"
     CDS             1095. .1706
                     /note="GTP-binding protein (SAS1)"
                     /codon_start=1
                     /protein_id="AAA33248.1"
                     /db_xref="GI:167875"
                     /translation="MTSPATNKPAAYDFLVKLLLIGDSGVGKSCLLLRFSDGSFTPSF
                     IATIGIDFKIRTIELEGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDEKS
                     FGSIRNWIRNIEQHASDSVNKMLIGNKCDMTEKKVVDSSRGKSLADEYGIKFLETSAK
                     NSVNVEEAFIGLAKDIKKRMIDTPNDPDHTICITPNNKKNTCC"
BASE COUNT      929 a    249 c    238 g    727 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.99e-43       Length:         2145
Score:                  517.50         Matches:        97
Percent Similarity:     71.36%         Conservative:   50
Best Local Similarity:  47.09%         Mismatches:     54
Query Match:            45.00%         Indels:         5
DB:                     3              Gaps:           3

US-09-817-199B-2 (1-223) x DDISAS1A (1-2145)

Qy         15 AlaProGluArgSerProProCysSerProSerTyrAspLeuThrGlyLysValMetLeu 34
              :::|||      :::   |||         :::||||||         |||::::::|||
Db       1101 TCTCCAGCAACAAATAAACCAGCA------GCCTACGATTTTTTAGTTAAATTACTTTTA 1154

Qy         35 LeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPhe 54
              :::||||||:::|||||||||||||:::|||   |||:::::::|||   ||||||:::|||
Db       1155 ATTGGGGATAGTGGTGTAGGAAAGTCATGTCTTTTATTACGTTTTTCTGATGGTTCTTTC 1214

Qy         55 LeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrVal 74
                 :::   :::||||||||||||||:::||||||||||||||:::   :::   :::   :::
Db       1215 ---ACACCAAGTTTCATCGCTACTATTGGTATCGATTTCAAAATTCGTACAATTGAATTA 1271

Qy         75 AspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSer 94
              :::|||   |||:::|||||||||||||||||||||||||||||||||||||||||:::
Db       1272 GAGGGTAAAAGAATTAAATTACAAATTTGGGACACTGCAGGTCAAGAAAGATTCAGAACT 1331

Qy         95 ValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsn 114
              :::|||   ||||||||||||   |||      :::||||||:::||||||:::|||:::
Db       1332 ATCACTACAGCATACTATCGCGGTGCTATGGGTATCCTATTGGTTTATGATGTCACTGAT 1391

Qy        115 LysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArg 134
              :::   ||||||   :::|||||||   |||:::      |||   :::::::|||
Db       1392 GAAAAATCTTTTGGTAGCATTAGAAATTGGATTAGAAATATCGAGCAACATGCTTCAGAC 1451

Qy        135 AspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArg 154
                 |||      ||||||:::|||||||||   ||||||:::   ::::::|||:::
Db       1452 TCAGTTAATAAAATGTTAATCGGTAATAAATGTGATATGACCGAAAAGAAAGTTGTTGAT 1511

Qy        155 SerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAla 174
              |||      |||::::::||||||   ||||||||||:::   ||||||||||||||||||
Db       1512 AGCTCAAGAGGTAAATCACTTGCAGACGAATATGGTATTAAATTTTTAGAAACTTCTGCC 1571

Qy        175 LysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg 194
              |||      :::||||||||||   |||||||:::   :::|||||||:::::|||   |||
Db       1572 AAAAACAGTGTAAATGTAGAGGAAGCCTTTATTGGTTTAGCAAAAGATATTAAAAAACGT 1631

Qy        195 AlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLys 214
                             :::::::|||         |||        :::
Db       1632 ATGATTGATACACCAAATGATCCTGATCATACCATA------TGCATTACTCCAAACAAT 1685

Qy        215 LysArgSerSerCysCys 220
              |||::::::::::||||||
Db       1686 AAGAAAAATACTTGTTGT 1703


Search completed: June 23, 2003, 20:33:25
Job time : 2493 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:48:33 ; Search time 49 Seconds
(without alignments)
492.451 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	223	10	US-09-817-199A-2	Sequence 2, Appli
2	1144	99.5	226	9	US-09-764-868-684	Sequence 684, App
3	1133	98.5	222	9	US-09-764-868-1106	Sequence 1106, Ap
4	1081	94.0	223	10	US-09-817-199A-4	Sequence 4, Appli
5	977	85.0	191	10	US-09-794-257-14	Sequence 14, Appl
6	977	85.0	191	12	US-10-051-986-3	Sequence 3, Appli
7	547	47.6	106	10	US-09-867-550-1812	Sequence 1812, Ap
8	535	46.5	139	9	US-09-764-868-688	Sequence 688, App
9	501	43.6	207	10	US-09-794-257-8	Sequence 8, Appli
10	500	43.5	190	10	US-09-822-860-5	Sequence 5, Appli
11	498	43.3	218	10	US-09-925-300-1571	Sequence 1571, Ap
12	497.5	43.3	162	10	US-09-834-765-766	Sequence 766, App
13	483	42.0	246	10	US-09-925-302-534	Sequence 534, App
14	472	41.0	201	10	US-09-967-736-8	Sequence 8, Appli
15	469.5	40.8	224	9	US-10-102-806-466	Sequence 466, App
16	468	40.7	201	10	US-09-967-736-3	Sequence 3, Appli
17	466	40.5	198	10	US-09-794-257-16	Sequence 16, Appl
18	466	40.5	198	10	US-09-945-173-5	Sequence 5, Appli
19	466	40.5	198	10	US-09-972-529-4	Sequence 4, Appli
20	458.5	39.9	222	10	US-09-820-003A-4	Sequence 4, Appli
21	448.5	39.0	212	10	US-09-350-874-67	Sequence 67, Appl
22	435	37.8	401	9	US-09-764-868-701	Sequence 701, App
23	429.5	37.3	188	9	US-09-764-868-1120	Sequence 1120, Ap
24	429	37.3	212	10	US-09-817-198A-2	Sequence 2, Appli
25	427.5	37.2	212	10	US-09-817-198A-4	Sequence 4, Appli
26	416	36.2	218	10	US-09-817-198A-5	Sequence 5, Appli
27	413.5	36.0	307	9	US-09-764-868-1100	Sequence 1100, Ap
28	413.5	36.0	312	10	US-09-925-302-783	Sequence 783, App
29	386	33.6	832	10	US-09-834-765-2	Sequence 2, Appli
30	377.5	32.8	222	9	US-09-764-868-1112	Sequence 1112, Ap
31	377.5	32.8	225	9	US-09-764-868-692	Sequence 692, App
32	373.5	32.5	213	10	US-09-794-257-5	Sequence 5, Appli
33	365	31.7	208	9	US-10-108-605-45	Sequence 45, Appl
34	364.5	31.7	624	10	US-09-834-765-5	Sequence 5, Appli
35	364.5	31.7	625	10	US-09-834-765-762	Sequence 762, App
36	360.5	31.3	168	10	US-09-834-765-765	Sequence 765, App
37	360	31.3	213	9	US-10-036-542-64	Sequence 64, Appl
38	360	31.3	217	9	US-10-036-542-89	Sequence 89, Appl
39	360	31.3	239	10	US-09-925-301-1077	Sequence 1077, Ap
40	359	31.2	217	10	US-09-988-974-3	Sequence 3, Appli
41	358.5	31.2	213	10	US-09-988-974-8	Sequence 8, Appli
42	352.5	30.7	217	10	US-09-925-300-1364	Sequence 1364, Ap
43	350	30.4	201	10	US-09-822-860-2	Sequence 2, Appli
44	349.5	30.4	161	10	US-09-834-765-763	Sequence 763, App
45	344.5	30.0	211	12	US-10-051-986-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-817-199A-2
; Sequence 2, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001187
; CURRENT APPLICATION NUMBER: US/09/817,199A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Human
US-09-817-199A-2

Query Match 100.0%; Score 1150; DB 10; Length 223;
Best Local Similarity 100.0%; Pred. No. 2e-108;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
```

RESULT 2

```
US-09-764-868-684
; Sequence 684, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 684
;   LENGTH: 226
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-868-684

  Query Match             99.5%;  Score 1144;  DB 9;  Length 226;
  Best Local Similarity   99.6%;  Pred. No. 8.1e-108;
  Matches  222;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 63

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ATVGIDFSNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 123

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 183

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        184 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 226


RESULT 3
US-09-764-868-1106
; Sequence 1106, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1106
;   LENGTH: 222
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (67)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (210)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1106

  Query Match             98.5%;  Score 1133;  DB 9;  Length 222;
  Best Local Similarity   99.1%;  Pred. No. 1e-106;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIA 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIA 60

Qy         62 TVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNI 121
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TVGIDFXNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNI 120

Qy        122 RAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVE 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVE 180

Qy        182 LAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||| ||||||||||||
Db        181 LAFLAIAKELKYRAGHQADEPSFQIRDYVXSQKKRSSCCSFM 222


RESULT 4
US-09-817-199A-4
; Sequence 4, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001187
;  CURRENT APPLICATION NUMBER: US/09/817,199A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-817-199A-4

  Query Match             94.0%;  Score 1081;  DB 10;  Length 223;
  Best Local Similarity   93.7%;  Pred. No. 1.9e-101;
  Matches  209;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              |||||||    |||||||||| ||:|||||||||||:|||||||||||||||||||||||
Db          1 MTGTPGAATAGDGEAPERSPPFSPNYDLTGKVMLLGDSGVGKTCFLIQFKDGAFLSGTFI 60

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              |||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db         61 ATVGIDFRNKVVTVDGARVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNQSSFDN 120

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||:||||||||||||||||||||:||||||||||
Db        121 IRAWLTEIHEYAQRDVVIMLLGNKADVSSERVIRSEDGETLAREYGVPFMETSAKTGMNV 180

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||| | |||||||||||||||||:
Db        181 ELAFLAIAKELKYRAGRQPDEPSFQIRDYVESQKKRSSCCSFV 223


RESULT 5
US-09-794-257-14
; Sequence 14, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 191
;   TYPE: PRT
;   ORGANISM: homo sapiens
```

US-09-794-257-14

Query Match 85.0%; Score 977; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 5e-91;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              ||||||||||||
Db        181 QKKRSSCCSFM 191
```

RESULT 6
US-10-051-986-3
; Sequence 3, Application US/10051986
; Patent No. US20020146770A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Tang, Y. Tom
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; Batra, Sajeev
; Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/051,986
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UCMCL5T01
; CLONE: 1528559
; SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-10-051-986-3

Query Match 85.0%; Score 977; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 5e-91;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              ||||||||||||
Db        181 QKKRSSCCSFM 191
```

RESULT 7
US-09-867-550-1812
; Sequence 1812, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206Alel Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1812
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1812

Query Match 47.6%; Score 547; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.1e-48;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVI 138
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVI 106
```

RESULT 8
US-09-764-868-688
; Sequence 688, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17

```
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 688
;    LENGTH: 139
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-764-868-688

  Query Match             46.5%;  Score 535;  DB 9;  Length 139;
  Best Local Similarity   71.5%;  Pred. No. 1.7e-46;
  Matches   98;  Conservative   19;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         84 WDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGN 143
              |||||||||||||||||||| |||||||:|||:|||||:|||||||||||| || :|||||
Db          1 WDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGN 60

Qy        144 KADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPS 203
              | | : |||:: |||| ||:|||:||:|||||||:||:||| ||||||| |:     ||
Db         61 KVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPR 120

Qy        204 FQIRDYVESQKKRSSCC 220
              |:: |||: : : :|||
Db        121 FRLHDYVKREGRGASCC 137


RESULT 9
US-09-794-257-8
; Sequence 8, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;    LENGTH: 207
;    TYPE: PRT
;    ORGANISM: homo sapiens
US-09-794-257-8

  Query Match             43.6%;  Score 501;  DB 10;  Length 207;
  Best Local Similarity   49.2%;  Pred. No. 8.2e-43;
  Matches   98;  Conservative   40;  Mismatches   59;  Indels    2;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| ::||||
Db          4 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTIELDGKKIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||: |:  | |:|  ||  |:|||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQA-DEPS 203
              ||: :| :  | || || :||: |||||||:  ||| ||  :|:::  :   :  |  |
Db        123 CDMNDKRQVSKERGEKLAIDYGIKFLETSAKSSANVEEAFFTLARDIMTKLNRKMNDSNS 182

Qy        204 FQIRDYVESQKKRSSCCSF 222
                    |:  : ||   ||
Db        183 AGAGGPVKITENRSKKTSF 201


RESULT 10
US-09-822-860-5
; Sequence 5, Application US/09822860
; Patent No. US20020146795A1
; GENERAL INFORMATION:
;  APPLICANT: ZHU, Shiaoping et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001214
;  CURRENT APPLICATION NUMBER: US/09/822,860
;  CURRENT FILING DATE:  2001-04-02
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;    LENGTH: 190
;    TYPE: PRT
;    ORGANISM: Discopyge ommata
US-09-822-860-5

  Query Match             43.5%;  Score 500;  DB 10;  Length 190;
  Best Local Similarity   50.3%;  Pred. No. 9.2e-43;
  Matches   96;  Conservative   42;  Mismatches   51;  Indels    2;  Gaps    2;

Qy         27 DLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDT 86
              |   |::|:||:|||||| | :| : || : |||:|:||||: : | :|| ::|||||||
Db          1 DYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTVELDGKKIKLQIWDT 59

Qy         87 AGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKAD 146
              |||||||::| |||| |  :: :|||||: |||||: |:  | |:|  ||  |:|||| |
Db         60 AGQERFRTITTAYYRGAMGIMKVYDITNEKSFDNIKNWIRNIEEHASSDVERMILGNKCD 119

Qy        147 MSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQ- 205
              |: :| :  | || || :||: |||||||: :||| ||: :|:::  :   : :| | |
Db        120 MNEKRQVSKERGEKLAIDYGIKFLETSAKSSINVEEAFITLARDIMTKLNKKMNENSLQE 179

Qy        206 IRDYVESQKKR 216
                | ::|  |:
Db        180 AVDKLKSPPKK 190


RESULT 11
US-09-925-300-1571
; Sequence 1571, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1571
;    LENGTH: 218
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-925-300-1571

  Query Match             43.3%;  Score 498;  DB 10;  Length 218;
  Best Local Similarity   48.5%;  Pred. No. 1.8e-42;
  Matches   99;  Conservative   36;  Mismatches   63;  Indels    6;  Gaps    3;

Qy         20 PPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRV 79
              |    :|||  |::|:||:|||||| | :| | || : |||:|:||||: | | : | ::
Db         18 PMAKKTYDLLFKLLLIGDSGVGKTCVLFRFSDDAF-NTTFISTIGIDFKIKTVELQGKKI 76

Qy         80 KLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIM 139
              |||||||||||||| ::| :||| |  ::|:|||||  ||:||  ||  | |:|  ||  |
Db         77 KLQIWDTAGQERFHTITTSYYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERM 136

Qy        140 LLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQA 199
              ||||| ||  :||:    || :|||:|: |.|||||  :|:| ||| :|:::  :
```

```
Db        137 LLGNKCDMDDKRVVPKGKGEQIAREHGIRFFETSAKANINIEKAFLTLAEDILRKT--PV 194

Qy        200 DEPSFQIRDYVES---QKKRSSCC 220
              ||: :  |          :| ||
Db        195 KEPNSENVDISSGGGVTGWKSKCC 218


RESULT 12
US-09-834-765-766
; Sequence 766, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
;  APPLICANT: Mary Faris
;  APPLICANT:  Pia M. Challita-Eid
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Steve Chappell Mitchell
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Aya Jakobovits
;  TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
;  TITLE OF INVENTION:  AND DETECTION OF CANCER
;  FILE REFERENCE: 129.6USU1
;  CURRENT APPLICATION NUMBER: US/09/834,765
;  CURRENT FILING DATE:  2001-09-21
;  PRIOR APPLICATION NUMBER: 60/197,647
;  PRIOR FILING DATE: 2000-04-12
;  NUMBER OF SEQ ID NOS: 770
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 766
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-765-766

  Query Match            43.3%;  Score 497.5;  DB 10;  Length 162;
  Best Local Similarity   55.6%;  Pred. No. 1.3e-42;
  Matches   90;  Conservative   35;  Mismatches   36;  Indels    1;  Gaps    1;

Qy         31 KVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQE 90
              |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:||||||||||
Db          2 KLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIWDTAGQE 60

Qy         91 RFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSE 150
              |||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:|||| |:: :
Db         61 RFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNKCDVNDK 120

Qy        151 RVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELK 192
              | :  | || || :|| |:|||||  :||| ||  :|:::|
Db        121 RQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIK 162


RESULT 13
US-09-925-302-534
; Sequence 534, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 534
;   LENGTH: 246
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-534

  Query Match            42.0%;  Score 483;  DB 10;  Length 246;
  Best Local Similarity   44.4%;  Pred. No. 6.9e-41;
  Matches   95;  Conservative   50;  Mismatches   57;  Indels   12;  Gaps    4;

Qy         16 PERSPPCS---PSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVV 72
              |  :|| |    :||   |::|:||:|||||| :|:| :  | : |:|:|:|||: : |
Db         35 PPPAPPVSAMAKAYDHLFKLLLIGDSGVGKTCLIIRFAEDNF-NNTYISTIGIDFKIRTV 93

Qy         73 TVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYA 132
              ::|| ::|||:|||||||||:::| |||| |  ::|:||||:: ||:||: |:  | | |
Db         94 DIEGKKIKLQVWDTAGQERFKTITTAYYRGAMGIILVYDITDEKSFENIQNWMKSIKENA 153

Qy        133 QRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELK 192
                |  :||||| || ::| :: |  : ||||:|: | |||||: |||: || ::|:::
Db        154 SAGVERLLLGNKCDMEAKRKVQKEQADKLAREHGIRFFETSAKSSMNVDEAFSSLARDIL 213

Qy        193 Y-----RAGHQADEPSFQIRDYVESQKKRSSCCS 221
                  |:|:    ||    |    || :: ||
Db        214 LKSGGRRSGNGNKPPS---TDLKTCDKKNTNKCS 244


RESULT 14
US-09-967-736-8
; Sequence 8, Application US/09967736
; Patent No. US20020103340A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Lal, Preeti
;                   Corley, Neil C.
;                   Shah, Purvi
;        TITLE OF INVENTION: RAB PROTEINS
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/967,736
;             FILING DATE: 28-Sep-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/154,602
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0367 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-855-0555
;             TELEFAX: 415-845-4166
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 201 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GenBank
;             CLONE: 57006
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-967-736-8

  Query Match            41.0%;  Score 472;  DB 10;  Length 201;
  Best Local Similarity   43.1%;  Pred. No. 6.8e-40;
  Matches   87;  Conservative   51;  Mismatches   58;  Indels    6;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
```

```
              :| ||   |::|:||:||||:| |::| |  : : ::|:|:|:||: : : :||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDTY-TESYISTIGVDFKIRTIELDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||:|| :||| |  ::::||:|:: |: |:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTVTSSYYRGAHGIIVVYDVTDQESYANVKQWLQEIDRYASENVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQA--- 199
              ||:|:::::|: :   :  |  ||||||||||   ||| ||: :| |:| | |  |
Db        121 NKSDLTTKKVVDNTTAKEFADSLGVPFLETSAKNATNVEQAFMTMAAEIKKRMGPGAASG 180

Qy        200 -DEPSFQIRDYVESQKKRSSCC 220
               : |: :| |    :     ||
Db        181 GERPNLKI-DSTPVKSASGGCC 201


RESULT 15
US-10-102-806-466
; Sequence 466, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 466
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-102-806-466

  Query Match            40.8%;  Score 469.5;  DB 9;  Length 224;
  Best Local Similarity  40.2%;  Pred. No. 1.4e-39;
  Matches   90;  Conservative   56;  Mismatches   65;  Indels   13;  Gaps    5;

Qy          8 VATRDGE-----APERSPPC--SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              :  |: |     | ||:     :| ||   |::|:||:||||:| |::| |  : : ::|
Db          3 ILEREAEQSRLGATERAAAAAMNPEYDYLFKLLLIGDSGVGKSCLLLRFADDTY-TESYI 61

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              :|:|:||: : : :||  :|||||||||||||||::| :||| |  ::::||:|:: |: |
Db         62 STIGVDFKIRTIELDGKTIKLQIWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESYAN 121

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              :: || ||  ||  :|  :|:|||:|:::::|: :   :  |   |:|||||||   ||
Db        122 VKQWLQEIDRYASENVNKLLVGNKSDLTTKKVVDNTTAKEFADSLGIPFLETSAKNATNV 181

Qy        181 ELAFLAIAKELKYRAGHQA----DEPSFQIRDYVESQKKRSSCC 220
              | ||: :| |:| | |  |    : |: :| |    :     ||
Db        182 EQAFMTMAAEIKKRMGPGAASGGERPNLKI-DSTPVKPAGGGCC 224


Search completed: June 18, 2003, 15:56:57
Job time : 49 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:48:03 ; Search time 28 Seconds
(without alignments)
234.333 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	977	85.0	191	4	US-09-075-454-3	Sequence 3, Appli
2	745	64.8	190	2	US-08-824-873-3	Sequence 3, Appli
3	745	64.8	190	3	US-09-198-184-3	Sequence 3, Appli
4	666	57.9	190	2	US-08-824-873-1	Sequence 1, Appli
5	666	57.9	190	3	US-09-198-184-1	Sequence 1, Appli
6	513.5	44.7	207	2	US-08-824-873-4	Sequence 4, Appli
7	513.5	44.7	207	3	US-09-198-184-4	Sequence 4, Appli
8	504	43.8	205	2	US-08-531-525-25	Sequence 25, Appl
9	504	43.8	205	2	US-08-718-270A-25	Sequence 25, Appl
10	486	42.3	207	2	US-08-531-525-35	Sequence 35, Appl
11	486	42.3	207	2	US-08-718-270A-35	Sequence 35, Appl
12	478.5	41.6	198	2	US-08-531-525-51	Sequence 51, Appl
13	478.5	41.6	198	2	US-08-718-270A-51	Sequence 51, Appl
14	478.5	41.6	215	2	US-08-531-525-10	Sequence 10, Appl
15	478.5	41.6	215	2	US-08-718-270A-10	Sequence 10, Appl
16	472	41.0	201	2	US-08-916-901-8	Sequence 8, Appli
17	472	41.0	201	4	US-09-154-602-8	Sequence 8, Appli
18	471	41.0	202	2	US-08-531-525-14	Sequence 14, Appl
19	471	41.0	202	2	US-08-718-270A-14	Sequence 14, Appl
20	468	40.7	201	2	US-08-916-901-3	Sequence 3, Appli
21	468	40.7	201	4	US-09-154-602-3	Sequence 3, Appli
22	458	39.8	201	2	US-08-531-525-13	Sequence 13, Appl
23	458	39.8	201	2	US-08-718-270A-13	Sequence 13, Appl
24	450.5	39.2	194	2	US-08-531-525-34	Sequence 34, Appl
25	450.5	39.2	194	2	US-08-718-270A-34	Sequence 34, Appl
26	448.5	39.0	212	4	US-09-399-913-67	Sequence 67, Appl
27	446.5	38.8	212	2	US-08-531-525-18	Sequence 18, Appl
28	446.5	38.8	212	2	US-08-718-270A-18	Sequence 18, Appl
29	426.5	37.1	210	2	US-08-531-525-16	Sequence 16, Appl
30	426.5	37.1	210	2	US-08-718-270A-16	Sequence 16, Appl
31	410	35.7	208	2	US-08-531-525-17	Sequence 17, Appl
32	410	35.7	208	2	US-08-718-270A-17	Sequence 17, Appl
33	409.5	35.6	218	2	US-08-531-525-19	Sequence 19, Appl
34	409.5	35.6	218	2	US-08-718-270A-19	Sequence 19, Appl
35	401	34.9	88	3	US-08-665-259-12	Sequence 12, Appl
36	401	34.9	88	3	US-08-762-500-12	Sequence 12, Appl
37	391.5	34.0	203	2	US-08-766-551-8	Sequence 8, Appli
38	389.5	33.9	208	2	US-08-531-525-15	Sequence 15, Appl
39	389.5	33.9	208	2	US-08-718-270A-15	Sequence 15, Appl
40	381	33.1	214	2	US-08-531-525-52	Sequence 52, Appl
41	381	33.1	214	2	US-08-718-270A-52	Sequence 52, Appl
42	364	31.7	203	2	US-08-531-525-21	Sequence 21, Appl
43	364	31.7	203	2	US-08-718-270A-21	Sequence 21, Appl
44	361	31.4	213	2	US-08-531-525-11	Sequence 11, Appl
45	361	31.4	213	2	US-08-718-270A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-075-454-3
; Sequence 3, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UCMCL5T01

```
;      CLONE:  1528559
US-09-075-454-3

  Query Match             85.0%;  Score 977;  DB 4;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-101;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              |||||||||||
Db        181 QKKRSSCCSFM 191


RESULT 2
US-08-824-873-3
; Sequence 3, Application US/08824873
; Patent No. 5843717
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,873
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  619734
US-08-824-873-3

  Query Match             64.8%;  Score 745;  DB 2;  Length 190;
  Best Local Similarity   73.4%;  Pred. No. 3e-75;
  Matches  138;  Conservative   27;  Mismatches   23;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:||:|||||| |::|||||||:||||:||||||||||: |||::|||||||||||||
Db          1 MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDGMKVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||| |||||||||| |||||:|||||| ||||:|||:|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQDVVLMLLGNKVDSTQERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: |||| ||:|||:||:||||||:|:||:||| ||||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELKQRSTKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : :  |||
Db        181 EGRGVSCC 188


RESULT 3
US-09-198-184-3
; Sequence 3, Application US/09198184
; Patent No. 6010859
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/198,184
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/824,873
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  619734
US-09-198-184-3

  Query Match             64.8%;  Score 745;  DB 3;  Length 190;
  Best Local Similarity   73.4%;  Pred. No. 3e-75;
  Matches  138;  Conservative   27;  Mismatches   23;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:||:|||||| |::|||||||:||||:||||||||||: |||::|||||||||||||
Db          1 MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDGMKVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
```

```
                ||||||||||| |||||||||| ||||:|||||| ||||:|||:|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQDVVLMLLGNKVDSTQERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: |||| ||:|||:||:|||||:|:||:||| |||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELKQRSTKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : :  |||
Db        181 EGRGVSCC 188


RESULT 4
US-08-824-873-1
; Sequence 1, Application US/08824873
; Patent No. 5843717
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,873
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PANCNOT04
;      CLONE:  738957
US-08-824-873-1

  Query Match             57.9%;  Score 666;  DB 2;  Length 190;
  Best Local Similarity   67.0%;  Pred. No. 2e-66;
  Matches  126;  Conservative   25;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:||:|||||| |   :   | |       |||||||: |||||||||:|||||||||
Db          1 MLVGDSGVGKTCLLGAIQGWCFPGGDLHLHRSIDFRNKVLDVDGVKVKLQMWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||| |||||||:|||:|||||:|||||||||| || :|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGNKVDSAHERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: |||| ||:|||:||:||||||||:||:||| |||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : : :|||
Db        181 EGRGASCC 188


RESULT 5
US-09-198-184-1
; Sequence 1, Application US/09198184
; Patent No. 6010859
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/198,184
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/824,873
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PANCNOT04
;      CLONE:  738957
US-09-198-184-1

  Query Match             57.9%;  Score 666;  DB 3;  Length 190;
  Best Local Similarity   67.0%;  Pred. No. 2e-66;
  Matches  126;  Conservative   25;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:||:|||||| |   :   | |       |||||||: |||||||||:|||||||||
Db          1 MLVGDSGVGKTCLLGAIQGWCFPGGDLHLHRSIDFRNKVLDVDGVKVKLQMWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||| |||||||:|||:|||||:|||||||||| || :|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGNKVDSAHERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: |||| ||:|||:||:||||||||:||:||| |||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : : :|||
Db        181 EGRGASCC 188
```

```
RESULT 6
US-08-824-873-4
; Sequence 4, Application US/08824873
; Patent No. 5843717
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,873
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  207 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  234746
US-08-824-873-4

  Query Match           44.7%;  Score 513.5;  DB 2;  Length 207;
  Best Local Similarity  50.0%;  Pred. No. 2.4e-49;
  Matches   99;  Conservative   41;  Mismatches   53;  Indels    5;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: ||||||  |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        183 QGSNQGVKITPDQQKRSS 200


RESULT 7
US-09-198-184-4
; Sequence 4, Application US/09198184
; Patent No. 6010859
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/198,184
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/824,873
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  207 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  234746
US-09-198-184-4

  Query Match           44.7%;  Score 513.5;  DB 3;  Length 207;
  Best Local Similarity  50.0%;  Pred. No. 2.4e-49;
  Matches   99;  Conservative   41;  Mismatches   53;  Indels    5;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: ||||||  |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        183 QGSNQGVKITPDQQKRSS 200


RESULT 8
US-08-531-525-25
; Sequence 25, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
```

```
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  205 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Canis familiaris
US-08-531-525-25

  Query Match             43.8%;  Score 504;  DB 2;  Length 205;
  Best Local Similarity   50.0%;  Pred. No. 2.7e-48;
  Matches   99;  Conservative   41;  Mismatches   52;  Indels    6;  Gaps    3;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:|||: : : :|| |:|||||
Db          3 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         62 DTAGQERFRTITTAYYR-AMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 120

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        121 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 180

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        181 QGSNQGVKITPDQQKRSS 198


RESULT 9
US-08-718-270A-25
; Sequence 25, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/718,270A
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/531,525
;      FILING DATE:  21-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,091
;      FILING DATE:  21-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  78-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  205 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Canis familiaris
US-08-718-270A-25

  Query Match             43.8%;  Score 504;  DB 2;  Length 205;
  Best Local Similarity   50.0%;  Pred. No. 2.7e-48;
  Matches   99;  Conservative   41;  Mismatches   52;  Indels    6;  Gaps    3;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:|||: : : :|| |:|||||
Db          3 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         62 DTAGQERFRTITTAYYR-AMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 120

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        121 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 180

Qy        201 EPSFQIRDYVESQKKRSS 218
              : |||       |:||||
Db        181 QGSNQGVKITPDQQKRSS 198


RESULT 10
US-08-531-525-35
; Sequence 35, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  207 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Discopyge ommata
US-08-531-525-35

  Query Match           42.3%;  Score 486;  DB 2;  Length 207;
  Best Local Similarity   48.1%;  Pred. No. 2.8e-46;
  Matches  101;  Conservative   44;  Mismatches   49;  Indels   16;  Gaps    6;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : | :|| ::|||||
Db          3 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTVELDGKKIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  :: : ||||: |||||: |:  | |:|  ||  |:||||
Db         62 DTAGQERFRTIT-AYYRGAMGIMKV-DITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 119

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
               ||: :| :  | || || :||: |||||||: :||| ||: :|:::  :   : :| |
Db        120 CDMNEKRQVSKERGEKLAIDYGIKFLETSAKSSINVEEAFITLARDIMTKLNKKMNENSL 179

Qy        205 QIRDYVE---------SQKKR--SSCCSFM 223
              |  : |:         ||||:  |  || :
Db        180 Q--EAVDKLKSPPKKPSQKKKQLSFRCSLL 207


RESULT 11
US-08-718-270A-35
; Sequence 35, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/718,270A
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/531,525
;      FILING DATE:  21-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,091
;      FILING DATE:  21-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  78-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  207 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Discopyge ommata
US-08-718-270A-35

  Query Match           42.3%;  Score 486;  DB 2;  Length 207;
  Best Local Similarity   48.1%;  Pred. No. 2.8e-46;
  Matches  101;  Conservative   44;  Mismatches   49;  Indels   16;  Gaps    6;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : | :|| ::|||||
Db          3 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTVELDGKKIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  :: : ||||: |||||: |:  | |:|  ||  |:||||
Db         62 DTAGQERFRTIT-AYYRGAMGIMKV-DITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 119

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
               ||: :| :  | || || :||: |||||||: :||| ||: :|:::  :   : :| |
Db        120 CDMNEKRQVSKERGEKLAIDYGIKFLETSAKSSINVEEAFITLARDIMTKLNKKMNENSL 179

Qy        205 QIRDYVE---------SQKKR--SSCCSFM 223
              |  : |:         ||||:  |  || :
Db        180 Q--EAVDKLKSPPKKPSQKKKQLSFRCSLL 207


RESULT 12
US-08-531-525-51
; Sequence 51, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
```

```
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Canis familiaris
US-08-531-525-51

  Query Match             41.6%;  Score 478.5;  DB 2;  Length 198;
  Best Local Similarity   48.7%;  Pred. No. 1.8e-45;
  Matches   97;  Conservative   36;  Mismatches   59;  Indels    7;  Gaps    4;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :|||  |::|:||:|||||| | :| | || : ||| ::|||: | | : | ::|||||
Db          4 TYDLLFKLLLIGDSGVGKTCVLFRFSDDAF-NTTFI-SIGIDFKIKTVELQGKKIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||| ::| :||| |  ::|:|||||  ||:||  ||  | |:|  ||  ||||||
Db         62 DTAGQERFHTITTSYYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLLGNK 121

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
              ||  :||:    || :|||:|: | |||||  :|:| ||| :|::  :      ||:
Db        122 CDMDDKRVVPKGKGEQIAREHGIRFFETSAKVNINIEKAFLTLAEDILRKT--PVKEPNS 179

Qy        205 QIRDYVES---QKKRSSCC 220
              :  |          :| ||
Db        180 ENVDISSGGGVTGWKSKCC 198


RESULT 13
US-08-718-270A-51
; Sequence 51, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/718,270A
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/531,525
;      FILING DATE:  21-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,091
;      FILING DATE:  21-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  78-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Canis familiaris
US-08-718-270A-51

  Query Match             41.6%;  Score 478.5;  DB 2;  Length 198;
  Best Local Similarity   48.7%;  Pred. No. 1.8e-45;
  Matches   97;  Conservative   36;  Mismatches   59;  Indels    7;  Gaps    4;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :|||  |::|:||:|||||| | :| | || : ||| ::|||: | | : | ::|||||
Db          4 TYDLLFKLLLIGDSGVGKTCVLFRFSDDAF-NTTFI-SIGIDFKIKTVELQGKKIKLQIW 61

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||| ::| :||| |  ::|:|||||  ||:||  ||  | |:|  ||  ||||||
Db         62 DTAGQERFHTITTSYYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLLGNK 121

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
              ||  :||:    || :|||:|: | |||||  :|:| ||| :|::  :      ||:
Db        122 CDMDDKRVVPKGKGEQIAREHGIRFFETSAKVNINIEKAFLTLAEDILRKT--PVKEPNS 179

Qy        205 QIRDYVES---QKKRSSCC 220
              :  |          :| ||
Db        180 ENVDISSGGGVTGWKSKCC 198


RESULT 14
US-08-531-525-10
; Sequence 10, Application US/08531525
; Patent No. 5840683
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5840683le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptides Inhibiting the Oncogenic Action
;    TITLE OF INVENTION:  of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
```

```
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,525
;      FILING DATE:  21-SEP-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  37-94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
US-08-531-525-10

  Query Match             41.6%;  Score 478.5;  DB 2;  Length 215;
  Best Local Similarity   45.1%;  Pred. No. 2e-45;
  Matches   96;  Conservative   50;  Mismatches   54;  Indels   13;  Gaps    6;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:||||||||||||| |::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQER-RTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 120

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR- 194
               :|:|||||  |:| : :  |: || |||: | |||||| :|||  | :| :::| |
Db        121 NKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVEEVFFSIGRDIKQRL 180

Qy        195 --AGHQADEPSFQIRDYVES-----QKKRSSCC 220
                 :|:  : :|    ::       ::|:||
Db        181 SDTDSRAEPATIKISQTDQAAGAGQATQKSACC 213


RESULT 15
US-08-718-270A-10
; Sequence 10, Application US/08718270A
; Patent No. 5910478
;  GENERAL INFORMATION:
;    APPLICANT:  Hlavka, Joseph J.
;    APPLICANT:  Pincus, Matthew R.
;    APPLICANT:  No. 5910478le, John F.
;    APPLICANT:  Abajian, Henry B.
;    APPLICANT:  Kende, Andrew S.
;    TITLE OF INVENTION:  Peptidomimetics Inhibiting
;    TITLE OF INVENTION:  the Oncogenic Action of P21 Ras
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/718,270A
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/531,525
;      FILING DATE:  21-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,091
;      FILING DATE:  21-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  78-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
US-08-718-270A-10

  Query Match             41.6%;  Score 478.5;  DB 2;  Length 215;
  Best Local Similarity   45.1%;  Pred. No. 2e-45;
  Matches   96;  Conservative   50;  Mismatches   54;  Indels   13;  Gaps    6;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:||||||||||||| |::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQER-RTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 120

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR- 194
               :|:|||||  |:| : :  |: || |||: | |||||| :|||  | :| :::| |
Db        121 NKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVEEVFFSIGRDIKQRL 180

Qy        195 --AGHQADEPSFQIRDYVES-----QKKRSSCC 220
                 :|:  : :|    ::       ::|:||
Db        181 SDTDSRAEPATIKISQTDQAAGAGQATQKSACC 213


Search completed: June 18, 2003, 15:56:01
Job time : 29 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:36:32 ; Search time 87 Seconds
(without alignments)
528.144 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	697	60.6	666	5	Q9VP48	Q9vp48 drosophila
2	648.5	56.4	388	5	Q95R32	Q95r32 drosophila
3	533	46.3	196	5	O02046	O02046 caenorhabdi
4	521.5	45.3	207	11	Q8VCF6	Q8vcf6 mus musculu
5	514.5	44.7	202	10	Q9SXT5	Q9sxt5 cicer ariet
6	513.5	44.7	202	10	Q08153	Q08153 pisum sativ
7	510	44.3	216	10	O49844	O49844 daucus caro
8	509.5	44.3	214	10	Q40218	Q40218 lotus japon
9	509	44.3	203	10	Q949E2	Q949e2 oryza sativ
10	509	44.3	216	10	Q40215	Q40215 lotus japon
11	509	44.3	218	10	Q9SF91	Q9sf91 arabidopsis
12	508.5	44.2	204	5	O15971	O15971 drosophila
13	507.5	44.1	202	10	Q40203	Q40203 lotus japon
14	507	44.1	216	10	Q9LZD4	Q9lzd4 arabidopsis
15	507	44.1	216	10	Q8VWF9	Q8vwf9 nicotiana t
16	506	44.0	203	10	Q9M7P5	Q9m7p5 capsicum an
17	506	44.0	216	10	Q40177	Q40177 lycopersico
18	505.5	44.0	201	5	Q94148	Q94148 caenorhabdi
19	505	43.9	203	10	O24112	O24112 nicotiana p
20	505	43.9	216	10	Q8W3J4	Q8w3j4 nicotiana t
21	504	43.8	216	10	Q8W3J3	Q8w3j3 nicotiana t
22	503	43.7	203	10	Q41340	Q41340 lycopersico
23	502	43.7	216	10	Q9FJF1	Q9fjf1 arabidopsis
24	501.5	43.6	203	10	Q8RU63	Q8ru63 oryza sativ
25	501	43.6	203	10	Q940Z7	Q940z7 arabidopsis
26	500	43.5	216	10	O24466	O24466 arabidopsis
27	499.5	43.4	216	10	Q8W3J2	Q8w3j2 nicotiana t
28	499	43.4	212	10	Q40217	Q40217 lotus japon
29	498	43.3	203	10	Q41338	Q41338 lycopersico
30	498	43.3	216	10	Q9SWV8	Q9swv8 lycopersico
31	497.5	43.3	202	10	Q9FPJ4	Q9fpj4 arabidopsis
32	497.5	43.3	202	10	Q08155	Q08155 pisum sativ
33	497.5	43.3	215	10	Q41023	Q41023 pisum sativ
34	496.5	43.2	202	10	Q9SEH3	Q9seh3 arabidopsis
35	495	43.0	215	10	Q40219	Q40219 lotus japon
36	495	43.0	216	10	Q41024	Q41024 pisum sativ
37	490.5	42.7	215	10	Q41022	Q41022 pisum sativ
38	489	42.5	222	10	Q96362	Q96362 brassica ca
39	488	42.4	203	10	Q08154	Q08154 pisum sativ
40	488	42.4	203	10	Q40569	Q40569 nicotiana t
41	487.5	42.4	202	10	Q8W4S8	Q8w4s8 arabidopsis
42	487.5	42.4	207	5	O18338	O18338 drosophila
43	487	42.3	200	4	Q9H0T3	Q9h0t3 homo sapien
44	486	42.3	202	10	Q9ZRH6	Q9zrh6 petunia hyb
45	484.5	42.1	215	10	Q41061	Q41061 pisum sativ

ALIGNMENTS

RESULT 1
Q9VP48
ID Q9VP48 PRELIMINARY; PRT; 666 AA.
AC Q9VP48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG7605 protein.
GN CG7605.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AE003594; AAF51708.1; -.
DR   HSSP; P05713; 3RAB.
DR   FlyBase; FBgn0037072; CG7605.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 2.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   666 AA;  72746 MW;  15355E368B316AD5 CRC64;

  Query Match             60.6%;  Score 697;  DB 5;  Length 666;
  Best Local Similarity   59.8%;  Pred. No. 1.5e-55;
  Matches  137;  Conservative   33;  Mismatches   49;  Indels   10;  Gaps    3;

Qy          3 GTPGAVAT---RDGEA------PERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGA 53
              |: || ||    | |        ::|     :|: |||::|||:|||||  ||:|:||
Db        433 GSEGASATLCKNAGRALIRMISSSKAPEEEDEFDIMGKVIMLGDSGVGKTSLLIRFRDGR 492

Qy         54 FLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDIT 113
              ::   |::||||||||||| ||| ||||||||||||||||||||||| |||||||:|
Db        493 YVPSYFLSTVGIDFRNKVVVVDGTRVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDVT 552

Qy        114 NKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMS-SERVIRSEDGETLAREYGVPFLET 172
              ||:::|||||| || |||| ||||:|:|||||  | ||| :: |||| | ||: |||:||
Db        553 NKTTYDNIRAWLGEIREYAQEDVVIVLIGNKADCSGSERQVKREDGERLGREHNVPFMET 612

Qy        173 SAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCS 221
              |||||:||||:| |:|::|| |     |:  | : |:|    |  | |:
Db        613 SAKTGLNVELSFTAVARQLKSRGYEHGDDGKFNVHDFVRDNTKARSVCA 661


RESULT 2
Q95R32
ID   Q95R32      PRELIMINARY;      PRT;   388 AA.
AC   Q95R32;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   GH21984p.
GN   CG7605.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY061826; AAL27637.1; -.
DR   FlyBase; FBgn0037072; CG7605.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding.
SQ   SEQUENCE   388 AA;  43065 MW;  D57C5244CA53492F CRC64;

  Query Match             56.4%;  Score 648.5;  DB 5;  Length 388;
  Best Local Similarity   65.1%;  Pred. No. 2e-51;
  Matches  125;  Conservative   25;  Mismatches   41;  Indels    1;  Gaps    1;

Qy         31 KVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQE 90
              | :||||:||||| ||::::  | |  |:| |||||   |||| ||| ||||||||||||
Db        192 KTILLGDSGVGKTSFLVKYNTGEFRLGSFSATVGIALTNKVVVVDGTRVKLQIWDTAGQE 251

Qy         91 RFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMS-S 149
              |||||||||||| |||||||:|||:::|||||| || |||| ||||:|:||||| | |
Db        252 RFRSVTHAYYRDAHALLLLYDVTNKTTYDNIRAWLGEIREYAQEDVVIVLIGNKADCSGS 311

Qy        150 ERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDY 209
              || :: |||| | ||: |||:|||||||:||||:| |:|::|| |     |:  | : |:
Db        312 ERQVKREDGERLGREHNVPFMETSAKTGLNVELSFTAVARQLKSRGYEHGDDGRFNVHDF 371

Qy        210 VESQKKRSSCCS 221
              |    |  | |:
Db        372 VRDNTKARSVCA 383


RESULT 3
O02046
ID   O02046      PRELIMINARY;      PRT;   196 AA.
AC   O02046;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   W01H2.3 protein.
GN   W01H2.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Minx P., Wohldmann P.;
RT   "The sequence of C. elegans cosmid W01H2.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
```

```
RA   Waterston R.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AF000192; AAB52888.1; -.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   196 AA;  22072 MW;  B51059FBDB5C44B6 CRC64;

  Query Match           46.3%;  Score 533;  DB 5;  Length 196;
  Best Local Similarity   54.5%;  Pred. No. 3.3e-41;
  Matches  109;  Conservative   34;  Mismatches   41;  Indels   16;  Gaps    5;

Qy         31 KVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQE 90
              |||||||:  |||| ||::|||||||:  ||:|||||:|||::|:   :|||||||||||
Db          4 KVMLLGDSCTGKTCLLIRYKDGAFLNNNFISTVGIDYRNKLITMGDKKVKLQIWDTAGQE 63

Qy         91 RFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSE 150
              |||||| :||||| ||||:||| |::||:| | ||::| || :  | : |:||| |:
Db         64 RFRSVTTSYYRDADALLLVYDIANRASFENCRNWLSQIKEYGKEAVQVTLVGNKCDL--P 121

Qy        151 RVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAK---ELKY---RAGHQADEPSF 204
              | : :::|: ||  | :||:||||||| ||: ||| :|:   :|||     |  ||  |
Db        122 RAVPTDEGKRLAEAYQIPFMETSAKTGFNVDRAFLGLAERMLKLKYGFVPGGEMADTIS- 180

Qy        205 QIRDYVESQKKR--SSCCSF 222
                  |   ||   : ||:|.
Db        181 -----VADTKKPEIARCCTF 195


RESULT 4
Q8VCF6
ID   Q8VCF6      PRELIMINARY;      PRT;   207 AA.
AC   Q8VCF6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to mel transforming oncogene (derived from cell line
DE   NK14)-RAB8 homolog.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC019990; AAH19990.1; -.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
SQ   SEQUENCE   207 AA;  23668 MW;  AC89DC85588FB8F8 CRC64;

  Query Match           45.3%;  Score 521.5;  DB 11;  Length 207;
  Best Local Similarity   49.3%;  Pred. No. 4.1e-40;
  Matches  101;  Conservative   42;  Mismatches   55;  Indels    7;  Gaps    3;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR-----AGHQA 199
              |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :      |:
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        200 DEPSFQIRDYVESQKKRSSC-CSFM 223
                 |  ::  || ||: |   || :
Db        183 QGSSHGVKITVEQQKRTSFFRCSLL 207


RESULT 5
Q9SXT5
ID   Q9SXT5      PRELIMINARY;      PRT;   202 AA.
AC   Q9SXT5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Rab-type small GTP-binding protein.
OS   Cicer arietinum (Chickpea) (Garbanzo).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX   NCBI_TaxID=3827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ILC3279; TISSUE=LEAF;
RA   Ichinose Y., Tiemann K., Schwenger-Erger C., Toyoda K., Hein F.,
RA   Hanselle T., Barz W.;
RT   "Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and
RT   Elicited Cell Cultures as Detected by Differential cDNA-
RT   Hybridization.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AB024994; BAA76422.1; -.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   202 AA;  22487 MW;  8CC4FCBAAC0C8CCE CRC64;

  Query Match           44.7%;  Score 514.5;  DB 10;  Length 202;
  Best Local Similarity   47.3%;  Pred. No. 1.7e-39;
  Matches   96;  Conservative   47;  Mismatches   53;  Indels    7;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              :| ||   |::|:||:||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQWLNEIDRYASENVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
              || |::: :|: ||  :  | | |:||:|||||   ||| ||:|:| |:| |   |
Db        121 NKCDLAANKVVSSETAKAFADEIGIPFMETSAKNATNVEQAFMAMAAEIKNRMASQPANN 180

Qy        199 ADEPSFQIRDYVESQKKRSSCCS 221
              |  |: |||    :||  | |||
Db        181 ARPPTVQIRGQPMNQK--SGCCS 201


RESULT 6
Q08153
```

```
ID   Q08153      PRELIMINARY;      PRT;   202 AA.
AC   Q08153;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   GTP-binding protein.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RX   MEDLINE=94290824; PubMed=8019783;
RA   Nagano Y., Murai N., Matsuno R., Sasaki Y.;
RT   "Isolation and characterization of cDNAs that encode eleven small GTP-
RT   binding protein from Pisum sativum.";
RL   Plant Cell Physiol. 34:447-455(1993).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; D12548; BAA02116.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   202 AA;  22498 MW;  1C9304144E5D95B3 CRC64;

  Query Match             44.7%;  Score 513.5;  DB 10;  Length 202;
  Best Local Similarity   47.3%;  Pred. No. 2.1e-39;
  Matches   96;  Conservative   47;  Mismatches   53;  Indels    7;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              :| ||   |::|:||:||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              ||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQWLNEIDRYASENVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
              ||:|::  :|: ||  :  | | |:||:|||||   ||| ||:|:| |:| |   |
Db        121 NKSDLADNKVVSSETAKAFADEIGIPFMETSAKNANNVEQAFMAMAAEIKNRMASQPSNN 180

Qy        199 ADEPSFQIRDYVESQKKRSSCCS 221
              |  |: |||    :||  | |||
Db        181 ARPPTVQIRGQPVNQK--SGCCS 201


RESULT 7
O49844
ID   O49844      PRELIMINARY;      PRT;   216 AA.
AC   O49844;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Small GTP-binding protein.
GN   DC-RAB8.
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LUNGA DI AMSTERDAM;
RA   Albani D., Bernacchia G.;
RT   "Isolation of a cDNA encoding a rab-related small GTP-binding protein
RT   (Accession No. AJ001367) from carrot cells (PGR97-185).";
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AJ001367; CAA04701.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   216 AA;  23955 MW;  3277D1EEDE77175B CRC64;

  Query Match             44.3%;  Score 510;  DB 10;  Length 216;
  Best Local Similarity   46.5%;  Pred. No. 5e-39;
  Matches   99;  Conservative   47;  Mismatches   55;  Indels   12;  Gaps    5;

Qy         19 SPPCSP--SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||  |   ||   |::|:||:||||:| |::| ||:| : :|| |:|||: : : :||
Db          3 APPARPRADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIEMDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
               :|:|||||  |:| : :  |: || |||: | | |||| |||  | :|||::| |
Db        122 NKILVGNKADMDESKRAVPTSKGQALADEYGIKFFEASAKTNMNVEEVFFSIAKDIKQRL 181

Qy        196 GH---QADEPSFQIRDY-----VESQKKRSSCC 220
                    : :  : :|          :  ::|:||
Db        182 AETDSKTEPQTIKINQQEQGAGTSAASQKSACC 214


RESULT 8
Q40218
ID   Q40218      PRELIMINARY;      PRT;   214 AA.
AC   Q40218;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAB8D.
GN   RAB8D.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ROOT NODULES;
RX   MEDLINE=97231679; PubMed=9076991;
RA   Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT   "Identification of new protein species among 33 different small GTP-
RT   binding proteins encoded by cDNAs from Lotos japonicus, and expression
RT   of corresponding mRNAs in developing root nodules.";
RL   Plant J. 11:237-250(1997).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; Z73947; CAA98175.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
```

```
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   214 AA;  23627 MW;  5E1A6E83505E50D7 CRC64;

  Query Match             44.3%;  Score 509.5;  DB 10;  Length 214;
  Best Local Similarity   47.6%;  Pred. No. 5.4e-39;
  Matches  101;  Conservative   45;  Mismatches   55;  Indels   11;  Gaps    5;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              ||||||||||||||||::| |||| |  :||:||:|:::||:||| ||  | ::|  :|
Db         62 KRVKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDEASFNNIRNWLRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR- 194
               :|:||||||  |:| : :  |: || |||: | |||||| :|||  | :||:::| |
Db        122 NKILVGNKADMDESKRAVPTSKGQALADEYGIKFFETSAKTNLNVEEVFFSIARDIKQRL 181

Qy        195 --AGHQADEPSFQIRD----YVESQKKRSSCC 220
                |:|:  : :|           :||||
Db        182 ADTDHKAEPTTLKINQDSAAGAGEAANKSSCC 213


RESULT 9
Q949E2
ID   Q949E2      PRELIMINARY;      PRT;   203 AA.
AC   Q949E2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative GTP-binding protein.
GN   W455ERIPDK.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21329048; PubMed=11435398;
RA   Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.,
RA   Duesterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,
RA   Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,
RA   Bevan M., Bancroft I.;
RT   "Conservation of microstructure bewtween a sequenced region of the
RT   genome of rice and multiple segments of the genome of Arabidopsis
RT   thaliana.";
RL   Genome Res. 11:1167-1174(2001).
DR   EMBL; AJ307662; CAC39050.1; -.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding.
SQ   SEQUENCE   203 AA;  22504 MW;  B59DE515C09C5A82 CRC64;

  Query Match             44.3%;  Score 509;  DB 10;  Length 203;
  Best Local Similarity   46.6%;  Pred. No. 5.6e-39;
  Matches   95;  Conservative   49;  Mismatches   52;  Indels    8;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              :| ||   |::|:||:|||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              ||||||||||::| :||| |  ::::||:|::: ||:|:: || ||  || .:|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIIVYDVTDQESFNNVKQWLNEIDRYASDNVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
                ||:||::: :|: ||  :  | | |:||:|||||   ||| ||:|:|  :| |   |
Db        121 NKSDLTANKVVSSETAKAFADEMGIPFMETSAKNATNVEQAFMAMAASIKDRMASQPAAA 180

Qy        199 -ADEPSFQIRDYVESQKKRSSCCS 221
               |  |: |||    :||  :||||
Db        181 NAREPTVQIRGQPVNQK--TSCCS 202


RESULT 10
Q40215
ID   Q40215      PRELIMINARY;      PRT;   216 AA.
AC   Q40215;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAB8A.
GN   RAB8A.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ROOT NODULES;
RX   MEDLINE=97231679; PubMed=9076991;
RA   Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT   "Identification of new protein species among 33 different small GTP-
RT   binding proteins encoded by cDNAs from Lotos japonicus, and expression
RT   of corresponding mRNAs in developing root nodules.";
RL   Plant J. 11:237-250(1997).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; Z73944; CAA98172.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   216 AA;  23776 MW;  569926CCA8D1B489 CRC64;

  Query Match             44.3%;  Score 509;  DB 10;  Length 216;
  Best Local Similarity   45.2%;  Pred. No. 6.1e-39;
  Matches  103;  Conservative   46;  Mismatches   57;  Indels   22;  Gaps    5;

Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              | | | | |  |              ||   |::|:||:||||:| |::| ||:| : :||
Db          1 MAGAPAARARAD------------YDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFI 46

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
               |:||||: : : :|| |:|||||||||||||::| |||| |  :||:||:|:::||:|
Db         47 TTIGIDFKIRTIELDGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDEASFNN 106

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMN 179
              |: |:  | ::|  :|  :|:||||||  |:| : :  |: || |||: | |||||| :|
Db        107 IKNWIRNIEQHASDNVNKILVGNKADMDESKRAVPTSKGQALADEYGIKFFETSAKTNLN 166

Qy        180 VELAFLAIAKELKYR---AGHQADEPSFQIRDYVES----QKKRSSCC 220
              || ,| :||:::| |       :|:  : ||      |    |  : |||
Db        167 VEEVFFSIARDIKQRLADTDSRAEPQTIQINQPDASASGGQAAQKSCC 214


RESULT 11
Q9SF91
ID   Q9SF91      PRELIMINARY;      PRT;   218 AA.
AC   Q9SF91;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative Ras-like GTP-binding protein.
GN   F8A24.5.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA   Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA   Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F8A24 genomic sequence.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AC015985; AAF23246.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   218 AA;  24294 MW;  E4C701A49AE426A1 CRC64;

  Query Match           44.3%;  Score 509;  DB 10;  Length 218;
  Best Local Similarity  47.0%;  Pred. No. 6.2e-39;
  Matches  103;  Conservative   46;  Mismatches   56;  Indels   14;  Gaps    5;

Qy         15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTV 74
              || |:      ||   |::|:||:||||:| |::| |  | : :|| |:||||: : | :
Db          4 APARA---RSDYDYLIKLLLIGDSGVGKSCLLLRFSDDTFTT-SFITTIGIDFKIRTVEL 59

Qy         75 DGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR 134
              || |:||||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|
Db         60 DGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWMKNIEQHASD 119

Qy        135 DVVIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKY 193
               |  :|:|||||  |:| : :  |: || |||: | ||||||  |||  ||:|||::|
Db        120 SVNKILVGNKADMDESKRAVPTSKGQALADEYGIKFFETSAKTNQNVEQVFLSIAKDIKQ 179

Qy        194 R---AGHQADEPSFQI------RDYVESQKKRSSCCSFM 223
              |   :  :|:    :|       :    |  ::|:|||::
Db        180 RLTESDTKAEPQGIKITKQDANKASSSSTNEKSACCSYV 218


RESULT 12
O15971
ID   O15971      PRELIMINARY;      PRT;   204 AA.
AC   O15971;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAB10 protein (LD39986P).
GN   RAB10 OR DRAB10 OR CG17060.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON R; TISSUE=HEAD;
RX   MEDLINE=97228579; PubMed=9074639;
RA   Satoh A.K., Tokunaga F., Ozaki K.;
RT   "Rab proteins of Drosophila melanogaster: novel members of the Rab-
RT   protein family.";
RL   FEBS Lett. 404:65-69(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AE003572; AAF50924.1; -.
DR   EMBL; AB006189; BAA21744.1; -.
DR   EMBL; AY060425; AAL25464.1; -.
DR   HSSP; P05713; 3RAB.
DR   FlyBase; FBgn0015789; Rab10.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   204 AA;  23336 MW;  4E058761C6854920 CRC64;

  Query Match           44.2%;  Score 508.5;  DB 5;  Length 204;
  Best Local Similarity  49.8%;  Pred. No. 6.3e-39;
  Matches  100;  Conservative   42;  Mismatches   54;  Indels    5;  Gaps    3;
```

```
Qy          25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
               :|||  |::|:||:|||||| | :| | || | |||:|:||||: | | : | ::|||||
Db           5 TYDLLFKLLLIGDSGVGKTCILFRFSDDAFTS-TFISTIGIDFKIKTVELRGKKIKLQIW 63

Qy          85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
               |||||||| ::| :||| |  ::|:|||||: ||:||  ||  | |:|  ||  |:|||
Db          64 DTAGQERFHTITTSYYRGAMGIMLVYDITNEKSFENIVKWLRNIDEHANEDVEKMILGNK 123

Qy         145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKE-LKYRAGHQADEPS 203
                ||: :||:  | || :|||:|: |:|||||: :|:| ||  :|:  |   :| :: |
Db         124 CDMTDKRVVNKERGEAIAREHGIRFMETSAKSNINIERAFCELAEAILDKTSGRESAENQ 183

Qy         204 FQIRDYVESQKKR---SSCCS 221
                ::     :|:|    | ||:
Db         184 ERVIIDRRNQEKAPGYSKCCA 204


RESULT 13
Q40203
ID   Q40203      PRELIMINARY;      PRT;   202 AA.
AC   Q40203;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAB1C.
GN   RAB1C.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ROOT NODULES;
RX   MEDLINE=97231679; PubMed=9076991;
RA   Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT   "Identification of new protein species among 33 different small GTP-
RT   binding proteins encoded by cDNAs from Lotos japonicus, and expression
RT   of corresponding mRNAs in developing root nodules.";
RL   Plant J. 11:237-250(1997).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; Z73932; CAA98160.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   202 AA;  22633 MW;  91633EBFAD07662F CRC64;

  Query Match             44.1%;  Score 507.5;  DB 10;  Length 202;
  Best Local Similarity   46.8%;  Pred. No. 7.6e-39;
  Matches   95;  Conservative   46;  Mismatches   55;  Indels    7;  Gaps    3;

Qy          23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
               :| ||   |::|:||:||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db           2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy          83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
               |||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db          61 IWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQWLNEIDRYASENVNKLLVG 120

Qy         143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
               || |::  :|:  |  :  | | |:||:|||||   ||| ||:|:| |:| |   |
Db         121 NKCDLTENKVVSYETAKAFADEIGIPFMETSAKNATNVEQAFMAMAAEIKNRMASQPVNN 180

Qy         199 ADEPSFQIRDYVESQKKRSSCCS 221


                |  ||: |||    :||  | |||
Db         181 ARPETVQIRGQPVNQK--SGCCS 201


RESULT 14
Q9LZD4
ID   Q9LZD4      PRELIMINARY;      PRT;   216 AA.
AC   Q9LZD4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   GTP-binding protein-like (Putative GTP-binding protein).
GN   F12E4_300.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA   Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA   Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA   Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA   Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA   Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RT   "Full Length cDNA of gene F12E4_300 (GI:7378637).";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA   Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA   Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA   Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA   Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "Full Length cDNA of gene F12E4_300 (GI:7378637).";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR   EMBL; AL162751; CAB83313.1; -.
DR   EMBL; AY056351; AAL07200.1; -.
DR   EMBL; AY035124; AAK59629.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW   GTP-binding; Lipoprotein.
SQ   SEQUENCE   216 AA;  24037 MW;  50C7D3AF4DC86F67 CRC64;

  Query Match             44.1%;  Score 507;  DB 10;  Length 216;
  Best Local Similarity   47.5%;  Pred. No. 9.4e-39;
  Matches  103;  Conservative   46;  Mismatches   56;  Indels   12;  Gaps    4;

Qy          15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTV 74
               || |:      ||   |::|:||:||||:| |::| |  | : :|| |:||||: : | :
```

```
Db          4 APARA---RSDYDYLIKLLLIGDSGVGKSCLLLRFSDDTFTT-SFITTIGIDFKIRTVEL 59

Qy         75 DGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR 134
              || |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|
Db         60 DGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWMKNIEQHASD 119

Qy        135 DVVIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKY 193
              :|  :|:|||||  |:| : :  |: || |||: | |||||| :|||  |::|||::|
Db        120 NVNKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVENVFMSIAKDIKQ 179

Qy        194 R-------AGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |       |  |  : : |      |  ::|:|||::
Db        180 RLTETDTKAEPQGIKITKQDTAASSSTAEKSACCSYV 216


RESULT 15
Q8VWF9
ID   Q8VWF9      PRELIMINARY;      PRT;   216 AA.
AC   Q8VWF9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Ras-related protein RAB8-4 (Ras-related protein RAB8-2).
GN   RAB8-4 OR RAB8-2.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRIGHT YELLOW 2;
RA   Torimoto N., Shimada K., Ito K., Yamamoto K.;
RT   "Characterization of Rab8 from tabacco BY-2 cell.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB079023; BAB84325.1; -.
DR   EMBL; AB079021; BAB84323.1; -.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR002078; Sig54_interact.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
SQ   SEQUENCE   216 AA;  23971 MW;  D5E87FABD0D49E13 CRC64;

  Query Match             44.1%;  Score 507;  DB 10;  Length 216;
  Best Local Similarity   46.9%;  Pred. No. 9.4e-39;
  Matches  100;  Conservative   45;  Mismatches   56;  Indels   12;  Gaps    4;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
              :|:||||||  |:| : :  |: || |||: | |||||| ||||  | :||:::| |
Db        122 NKILVGNKADMDESKRAVPTSKGQALADEYGIKFFETSAKTNMNVEEVFFSIARDIKQRL 181

Qy        196 GHQADEPSFQIRDYVESQK--------KRSSCC 220
                 :   |     :| :         ::||||
Db        182 SESDSKTEPQAIRINQSDQAGTSGQAAQKSSCC 214


Search completed: June 18, 2003, 15:46:31
Job time : 87 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:42:32 ; Search time 24 Seconds
(without alignments)
385.384 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	223	1	RB37_HUMAN	Q96ax2 homo sapien
2	1081	94.0	223	1	RB37_MOUSE	Q9jkm7 mus musculu
3	753	65.5	190	1	RB26_HUMAN	Q9ulw5 homo sapien
4	745	64.8	190	1	RB26_RAT	P51156 rattus norv
5	517.5	45.0	203	1	SAS2_DICDI	P20791 dictyosteli
6	516	44.9	206	1	RAB8_MOUSE	P55258 mus musculu
7	513.5	44.7	207	1	RAB8_HUMAN	P24407 homo sapien
8	512.5	44.6	215	1	RAB1_BETVU	Q39433 beta vulgar
9	512	44.5	210	1	RAB8_DISOM	P22128 discopyge o
10	507	44.1	258	1	ARA5_ARATH	P28188 arabidopsis
11	502.5	43.7	207	1	RB8B_RAT	P70550 rattus norv
12	501	43.6	207	1	RB8B_HUMAN	Q92930 homo sapien
13	500	43.5	202	1	RIC1_ORYSA	P40392 oryza sativ
14	498	43.3	208	1	SAS1_DICDI	P20790 dictyosteli
15	495	43.0	203	1	YPT2_MAIZE	Q05737 zea mays (m
16	495	43.0	216	1	ARA3_ARATH	P28186 arabidopsis
17	494	43.0	200	1	RB10_CANFA	P24409 canis famil
18	494	43.0	200	1	RB10_HUMAN	O88386 homo sapien
19	490.5	42.7	200	1	RAO1_DISOM	P22127 discopyge o
20	487	42.3	200	1	YPT2_SCHPO	P17609 schizosacch
21	486	42.3	203	1	YPT1_CHLRE	Q39571 chlamydomon
22	479.5	41.7	208	1	YPT1_MAIZE	P16976 zea mays (m
23	478	41.6	203	1	YPT1_VOLCA	P31584 volvox cart
24	477.5	41.5	205	1	RAB1_LYMST	Q05974 lymnaea sta
25	476	41.4	217	1	YPT2_VOLCA	P36861 volvox cart
26	473.5	41.2	203	1	RB13_HUMAN	P51153 homo sapien
27	472	41.0	201	1	RB1B_RAT	P10536 rattus norv
28	465.5	40.5	202	1	RAB1_DISOM	P22125 discopyge o
29	464.5	40.4	203	1	RYL1_YARLI	P41924 yarrowia li
30	461	40.1	200	1	RB10_RAT	P35281 rattus norv
31	459.5	40.0	203	1	YPT1_NEUCR	P33723 neurospora
32	458.5	39.9	205	1	RB1A_HUMAN	P11476 homo sapien
33	456.5	39.7	205	1	RB1A_RAT	P05711 rattus norv
34	456	39.7	201	1	YPT1_PHYIN	Q01890 phytophthor
35	454	39.5	203	1	YPT1_SCHPO	P11620 schizosacch
36	450	39.1	215	1	SEC4_YEAST	P07560 saccharomyc
37	448.5	39.0	212	1	RAB2_HUMAN	P08886 homo sapien
38	448.5	39.0	212	1	RAB2_LYMST	Q05975 lymnaea sta
39	448.5	39.0	212	1	RAB2_MOUSE	P53994 mus musculu
40	448.5	39.0	212	1	RAB2_RABIT	Q01971 oryctolagus
41	448	39.0	213	1	YPT4_VOLCA	P36863 volvox cart
42	446.5	38.8	212	1	RAB2_RAT	P05712 rattus norv
43	440	38.3	213	1	YPT4_CHLRE	Q39570 chlamydomon
44	427.5	37.2	212	1	RB15_RAT	P35289 rattus norv
45	423.5	36.8	206	1	YPT1_YEAST	P01123 saccharomyc

ALIGNMENTS

RESULT 1
RB37_HUMAN
ID RB37_HUMAN STANDARD; PRT; 223 AA.
AC Q96AX2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-37.
GN RAB37.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Secretory granules (By similarity).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; BC016615; AAH16615.1; -.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 36 43 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 143 146 GTP (BY SIMILARITY).
FT DOMAIN 59 67 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 219 219 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 220 220 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 223 AA; 24815 MW; 5A7A4887BCBB84A8 CRC64;

Query Match 100.0%; Score 1150; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.4e-94;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
```

```
Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223


RESULT 2
RB37_MOUSE
ID   RB37_MOUSE     STANDARD;      PRT;   223 AA.
AC   Q9JKM7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ras-related protein Rab-37.
GN   RAB37.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20189834; PubMed=10722846;
RA   Masuda E.S., Luo Y., Young C., Shen M., Rossi A.B., Huang B.C., Yu S.,
RA   Bennett M.K., Payan D.G., Scheller R.H.;
RT   "Rab37 is a novel mast cell specific GTPase localized to secretory
RT   granules.";
RL   FEBS Lett. 470:61-64(2000).
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: Specifically expressed in the bone marrow mast
CC       cells.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF233582; AAF67162.1; -.
DR   HSSP; P01112; 1PLJ.
DR   MGD; MGI:1929945; Rab37.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT   NP_BIND      36     43       GTP (BY SIMILARITY).
FT   NP_BIND      85     89       GTP (BY SIMILARITY).
FT   NP_BIND     143    146       GTP (BY SIMILARITY).
FT   DOMAIN       59     67       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       219    219       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       220    220       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   223 AA;  24716 MW;  E72013621C12C605 CRC64;

  Query Match             94.0%;  Score 1081;  DB 1;  Length 223;
  Best Local Similarity   93.7%;  Pred. No. 2.9e-88;
  Matches  209;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              |||||||    |||||||||| ||:||||||||||||:||||||||||||||||||||||
Db          1 MTGTPGAATAGDGEAPERSPPFSPNYDLTGKVMLLGDSGVGKTCFLIQFKDGAFLSGTFI 60

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||||||||||| |||||||||||||||||||||||||||||||:|||||
Db         61 ATVGIDFRNKVVTVDGARVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNQSSFDN 120

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              |||||||||||||||||||||||||||:||||||||||||||||||||:||||||||||
Db        121 IRAWLTEIHEYAQRDVVIMLLGNKADVSSERVIRSEDGETLAREYGVPFMETSAKTGMNV 180

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||| | |||||||||||||||||||||:
Db        181 ELAFLAIAKELKYRAGRQPDEPSFQIRDYVESQKKRSSCCSFV 223


RESULT 3
RB26_HUMAN
ID   RB26_HUMAN     STANDARD;      PRT;   190 AA.
AC   Q9ULW5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ras-related protein Rab-26.
GN   RAB26.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=20496221; PubMed=11043516;
RA   Seki N., Yoshikawa T., Hattori A., Miyajima N., Muramatsu M.,
RA   Saito T.;
RT   "cDNA cloning of a human RAB26-related gene encoding a Ras-like
RT   GTP-binding protein on chromosome 16p13.3 region.";
RL   J. Hum. Genet. 45:309-314(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB027137; BAA84707.1; -.
DR   EMBL; BC007681; AAH07681.1; -.
DR   HSSP; P36017; 1EK0.
DR   Genew; HGNC:14259; RAB26.
DR   MIM; 605455; -.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
FT   NP_BIND       4     11       GTP (BY SIMILARITY).
FT   NP_BIND      53     57       GTP (BY SIMILARITY).
FT   NP_BIND     111    114       GTP (BY SIMILARITY).
FT   DOMAIN       27     35       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       187    187       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       188    188       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   190 AA;  21354 MW;  517387B7C0B3291C CRC64;

  Query Match             65.5%;  Score 753;  DB 1;  Length 190;
  Best Local Similarity   73.4%;  Pred. No. 2e-59;
  Matches  138;  Conservative   28;  Mismatches   22;  Indels    0;  Gaps    0;
```

```
Qy          33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
               ||:||:|||||| |::|||||||:||||:|||||||||: |||||||||:||||||||
Db           1 MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDGVKVKLQMWDTAGQERF 60

Qy          93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
               |||||||||| |||||||:|||:|||||:|||||||||| || :|||||| | : |||
Db          61 RSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGNKVDSAHERV 120

Qy         153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
               :: |||| ||:|||:||:|||||||:||:||| ||||||| |:     || |:: |||:
Db         121 VKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPRFRLHDYVKR 180

Qy         213 QKKRSSCC 220
               : : :|||
Db         181 EGRGASCC 188


RESULT 4
RB26_RAT
ID   RB26_RAT       STANDARD;      PRT;   190 AA.
AC   P51156;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Ras-related protein Rab-26.
GN   RAB26.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Pancreas;
RX   MEDLINE=95169156; PubMed=7864900;
RA   Wagner A.C.C., Strowski M.Z., Goeke B., Williams J.A.;
RT   "Molecular cloning of a new member of the Rab protein family, Rab 26,
RT   from rat pancreas.";
RL   Biochem. Biophys. Res. Commun. 207:950-956(1995).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PANCREAS, KIDNEY, BRAIN,
CC       SUBMANDIBULAR GLAND, AND LUNG.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U18771; AAA69955.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
FT   NP_BIND       4     11       GTP (BY SIMILARITY).
FT   NP_BIND      53     57       GTP (BY SIMILARITY).
FT   NP_BIND     111    114       GTP (BY SIMILARITY).
FT   DOMAIN       27     35       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       187    187       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       188    188       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   190 AA;  21457 MW;  96B2D1661F363031 CRC64;

  Query Match             64.8%;  Score 745;  DB 1;  Length 190;
  Best Local Similarity   73.4%;  Pred. No. 1e-58;
  Matches  138;  Conservative   27;  Mismatches   23;  Indels    0;  Gaps    0;

Qy          33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
               ||:||:|||||| |::|||||||:||||:|||||||||: |||::|||||||||||||||
Db           1 MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDGMKVKLQIWDTAGQERF 60

Qy          93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
               |||||||||| |||||||:|||:|||||:|||||||||| |||||:|||||| ||||:|||||| | : |||
Db          61 RSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQDVVLMLLGNKVDSTQERV 120

Qy         153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
               :: |||| ||:|||:||:|||||:|:||:||| ||||||| |:     || |:: |||:
Db         121 VKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELKQRSTKAPSEPRFRLHDYVKR 180

Qy         213 QKKRSSCC 220
               : :  |||
Db         181 EGRGVSCC 188


RESULT 5
SAS2_DICDI
ID   SAS2_DICDI     STANDARD;      PRT;   203 AA.
AC   P20791;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GTP-binding protein SAS2.
GN   SASB OR SAS2.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90220623; PubMed=2109188;
RA   Saxe S.A., Kimmel A.R.;
RT   "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
RT   with sequence similarities to essential genes in Saccharomyces
RT   cerevisiae.";
RL   Mol. Cell. Biol. 10:2367-2378(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89209367; PubMed=3149563;
RA   Saxe S.A., Kimmel A.R.;
RT   "Genes encoding novel GTP-binding proteins in Dictyostelium.";
RL   Dev. Genet. 9:259-265(1988).
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M34456; AAA33248.1; -.
DR   PIR; B34716; B34716.
DR   PIR; B61571; B61571.
DR   HSSP; P05713; 3RAB.
DR   DictyDb; DD05021; sasB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1..
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Palmitate.
FT   NP_BIND       22     29       GTP (BY SIMILARITY).
FT   NP_BIND       70     74       GTP (BY SIMILARITY).
FT   NP_BIND      128    131       GTP (BY SIMILARITY).
FT   LIPID        202    202       PALMITATE (BY SIMILARITY).
FT   LIPID        203    203       PALMITATE (BY SIMILARITY).
SQ   SEQUENCE   203 AA;  22642 MW;  33EFB2F451ECAC96 CRC64;
```

```
  Query Match             45.0%;  Score 517.5;  DB 1;  Length 203;
  Best Local Similarity   47.1%;  Pred. No. 1.3e-38;
  Matches   97;  Conservative   50;  Mismatches   54;  Indels    5;  Gaps    3;

Qy         15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTV 74
              :|  : |   :||   |::|:||:||||:| |::| ||:| : :||||:||||: : : :
Db          3 SPATNKPA--AYDFLVKLLLIGDSGVGKSCLLLRFSDGSF-TPSFIATIGIDFKIRTIEL 59

Qy         75 DGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR 134
              :| |:|||||||||||::| |||| |  :||:||:|:: || :|| |:  | ::|
Db         60 EGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDEKSFGSIRNWIRNIEQHASD 119

Qy        135 DVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR 194
               |  ||:||| ||: ::|: |  |::|| ||||: ||||||  :||| ||: :||::| |
Db        120 SVNKMLIGNKCDMTEKKVVDSSRGKSLADEYGIKFLETSAKNSVNVEEAFIGLAKDIKKR 179

Qy        195 AGHQADEPSFQIRDYVESQKKRSSCC 220
                    ::|   |   :    |:::||
Db        180 MIDTPNDPDHTI--CITPNNKKNTCC 203


RESULT 6
RAB8_MOUSE
ID   RAB8_MOUSE     STANDARD;      PRT;   206 AA.
AC   P55258;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Ras-related protein Rab-8 (Oncogene c-mel).
GN   RAB8 OR MEL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91360267; PubMed=1886711;
RA   Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
RA   Johnson K.J.;
RT   "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT   genes.";
RL   Oncogene 6:1347-1351(1991).
CC   -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC       NEUROTRANSMITTER RELEASE.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53270; AAB19682.1; -.
DR   HSSP; P05713; 3RAB.
DR   MGD; MGI:96960; Mel.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW   Proto-oncogene.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       203    203       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   206 AA;  23557 MW;  49D832725D662942 CRC64;

  Query Match             44.9%;  Score 516;  DB 1;  Length 206;
  Best Local Similarity   49.5%;  Pred. No. 1.7e-38;
  Matches  101;  Conservative   41;  Mismatches   56;  Indels    6;  Gaps    3;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ----AD 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :        |
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKNWKATAA 182

Qy        201 EPSFQIRDYVESQKKRSSC-CSFM 223
                |  ::  || ||: |   || :
Db        183 GSSHGVKITVEQQKRTSFFRCSLL 206


RESULT 7
RAB8_HUMAN
ID   RAB8_HUMAN     STANDARD;      PRT;   207 AA.
AC   P24407;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Ras-related protein Rab-8 (Oncogene c-mel).
GN   RAB8 OR MEL.
OS   Homo sapiens (Human), and
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606, 9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RX   MEDLINE=94124602; PubMed=8294494;
RA   Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
RA   Tavitian A., Louvard D.;
RT   "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT   polarized cells but colocalizes with the tight junction marker ZO-1
RT   in polarized epithelial cells.";
RL   J. Cell Biol. 124:101-115(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RX   MEDLINE=91360267; PubMed=1886711;
RA   Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
RA   Johnson K.J.;
RT   "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT   genes.";
RL   Oncogene 6:1347-1351(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human; TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=C.familiaris; STRAIN=Cocker spaniel;
RX   MEDLINE=91061765; PubMed=2123294;
RA   Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT   "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
RT   line.";
RL   Mol. Cell. Biol. 10:6578-6585(1990).
CC   -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC       NEUROTRANSMITTER RELEASE.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X56741; CAA40065.1; -.
DR   EMBL; S53268; AAB19681.1; -.
DR   EMBL; BC002977; AAH02977.1; -.
DR   EMBL; X56385; CAB56776.1; -.
DR   PIR; B36364; B36364.
DR   PIR; B49647; B49647.
DR   HSSP; P05713; 3RAB.
DR   Genew; HGNC:7007; MEL.
DR   MIM; 165040; -.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW   Proto-oncogene.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       204    204       GERANYL-GERANYL (BY SIMILARITY).
FT   CONFLICT    177    183       LEGNSPQ -> WKATAP (IN REF. 2).
SQ   SEQUENCE   207 AA;  23668 MW;  AA52DBF54A2CD056 CRC64;

  Query Match             44.7%;  Score 513.5;  DB 1;  Length 207;
  Best Local Similarity   50.0%;  Pred. No. 2.9e-38;
  Matches   99;  Conservative   41;  Mismatches   53;  Indels    5;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        183 QGSNQGVKITPDQQKRSS 200


RESULT 8
RAB1_BETVU
ID   RAB1_BETVU     STANDARD;      PRT;   215 AA.
AC   Q39433;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Ras-related protein RAB1BV.
GN   RAB1BV.
OS   Beta vulgaris (Sugar beet).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX   NCBI_TaxID=3555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. D100 KS 38080;
RX   MEDLINE=96307523; PubMed=8680960;
RA   Dallery E., Quief S., Ben Jilany K.E., Kerckaert J.-P., Hagege D.;
RT   "Molecular cloning and structural analysis of cDNAs that encode 3
RT   small GTP-binding proteins from sugar beet.";
RL   C. R. Acad. Sci., III, Sci. Vie 319:91-97(1996).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z49152; CAA89021.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
FT   NP_BIND      22     29       GTP (BY SIMILARITY).
FT   NP_BIND      70     74       GTP (BY SIMILARITY).
FT   NP_BIND     128    131       GTP (BY SIMILARITY).
FT   LIPID       212    212       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       213    213       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   215 AA;  23787 MW;  781CC5A17F29B8AA CRC64;

  Query Match             44.6%;  Score 512.5;  DB 1;  Length 215;
  Best Local Similarity   47.6%;  Pred. No. 3.7e-38;
  Matches  101;  Conservative   47;  Mismatches   53;  Indels   11;  Gaps    5;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
               :|:|||||  |:| : :  |: || |||: | ||||| :|||  | :||:::| |
Db        122 NKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVEEVFFSIARDIKQRL 181

Qy        196 G----HQADEPSFQIRDYVESQKK---RSSCC 220
                   |  :||  |:   :|  :   :|:||
Db        182 ADSDTRQEAQPSITIKPADQSGNQAAAKSACC 213


RESULT 9
RAB8_DISOM
ID   RAB8_DISOM     STANDARD;      PRT;   210 AA.
AC   P22128;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Ras-related protein Rab-8 (ORA2).
OS   Discopyge ommata (Electric ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Torpediniformes; Narcinoidei; Narcinidae; Discopyge.
OX   NCBI_TaxID=7785;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Electric lobe;
RX   MEDLINE=91115900; PubMed=1899244;
RA   Ngsee J.K., Elferink L.A., Scheller R.H.;
RT   "A family of ras-like GTP-binding proteins expressed in electromotor
RT   neurons.";
RL   J. Biol. Chem. 266:2675-2680(1991).
```

```
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38391; AAA49232.1; -.
DR   PIR; B38625; B38625.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       207    207       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   210 AA;  24164 MW;  0F73EDB0DB8B9EEA CRC64;

  Query Match             44.5%;  Score 512;  DB 1;  Length 210;
  Best Local Similarity   48.6%;  Pred. No. 4e-38;
  Matches  102;  Conservative   44;  Mismatches   50;  Indels   14;  Gaps    4;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : | :|| ::|||||
Db          4 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTVELDGKKIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  :: :|||||: |||||: |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMKVYDITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
              ||: :| :  | || || :||: |||||||: :||| ||: :|::: :   : :| |
Db        123 CDMNEKRQVSKERGEKLAIDYGIKFLETSAKSSINVEEAFITLARDIMTKLNKKMNENSL 182

Qy        205 QIRDYVE---------SQKKR--SSCCSFM 223
              |  : |:         ||||:  |  || :
Db        183 Q--EAVDKLKSPPKKPSQKKKQLSFRCSLL 210


RESULT 10
ARA5_ARATH
ID   ARA5_ARATH     STANDARD;      PRT;   258 AA.
AC   P28188;
DT   01-DEC-1992 (Rel. 24, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ras-related protein ARA-5.
GN   ARA-5 OR AT1G02130 OR T7I23.6.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
RN   [2]
RP   SEQUENCE OF 64-258 FROM N.A.
RC   STRAIN=Various strains; TISSUE=Leaf;
RX   MEDLINE=92084144; PubMed=1748311;
RA   Anai T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
RA   Matsui M.;
RT   "Isolation and analysis of cDNAs encoding small GTP-binding proteins
RT   of Arabidopsis thaliana.";
RL   Gene 108:259-264(1991).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U89959; AAC24370.1; -.
DR   EMBL; D01027; BAA00832.1; -.
DR   PIR; PS0279; PS0279.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Multigene family.
FT   NP_BIND      70     77       GTP (BY SIMILARITY).
FT   NP_BIND     118    122       GTP (BY SIMILARITY).
FT   NP_BIND     176    179       GTP (BY SIMILARITY).
FT   DOMAIN       92    100       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       255    255       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       256    256       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   258 AA;  29169 MW;  34D278A0C504638F CRC64;

  Query Match             44.1%;  Score 507;  DB 1;  Length 258;
  Best Local Similarity   45.5%;  Pred. No. 1.4e-37;
  Matches   95;  Conservative   50;  Mismatches   56;  Indels    8;  Gaps    3;

Qy         18 RSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGV 77
              ||   :| ||   |::|:||:||||:| |::| | :::  ::|:|:|:||: : |  ||
Db         52 RSKTMNPEYDYLFKLLLIGDSGVGKSCLLLRFSDDSYVE-SYISTIGVDFKIRTVEQDGK 110

Qy         78 RVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVV 137
              :|||||||||||||::| :||| |  ::::||:|:: ||:|:: ||:||  ||  :|
Db        111 TIKLQIWDTAGQERFRTITSSYYRGAHGIIIVYDVTDEESFNNVKQWLSEIDRYASDNVN 170

Qy        138 IMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR--- 194
              :|:|||:|::  | |  |  :  | | |:||:|||||   ||| ||:|::  :| |
Db        171 KLLVGNKSDLTENRAIPYETAKAFADEIGIPFMETSAKDATNVEQAFMAMSASIKERMAS 230

Qy        195 --AGHQADEPSFQIRDYVESQKKRSSCCS 221
                ||: |  |: |||    :||  : |||
Db        231 QPAGNNARPPTVQIRGQPVAQK--NGCCS 257
```

```
RESULT 11
RB8B_RAT
ID   RB8B_RAT       STANDARD;      PRT;   207 AA.
AC   P70550;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Ras-related protein Rab-8B.
GN   RAB8B.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96393028; PubMed=8799816;
RA   Armstrong J., Thompson N., Squire J.H., Smith J., Hayes B., Solari R.;
RT   "Identification of a novel member of the Rab8 family from the rat
RT   basophilic leukaemia cell line, RBL.2H3.";
RL   J. Cell Sci. 109:1265-1274(1996).
CC   -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC       NEUROTRANSMITTER RELEASE (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS OF EXPRESSION IN THE SPLEEN,
CC       TESTIS AND BRAIN.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U53475; AAA99782.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein; Protein transport.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       204    204       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   207 AA;  23603 MW;  4A41AB26BF9DCAF4 CRC64;

  Query Match             43.7%;  Score 502.5;  DB 1;  Length 207;
  Best Local Similarity   47.3%;  Pred. No. 2.7e-37;
  Matches   97;  Conservative   45;  Mismatches   56;  Indels    7;  Gaps    3;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| ::|||||
Db          4 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTIELDGKKIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||: |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSF 204
              ||: :| :  | || || :||: |||||||:  ||| ||  :|::: :  : :: :
Db        123 CDMNDKRQVSKERGEKLAIDYGIKFLETSAKSSTNVEEAFFTLARDIMTKLNRKMNDSNS 182

Qy        205 QIR----DYVESQKKRSSC--CSFM 223
                  ||: |::|   || :
Db        183 SGAGGPVKITESRSKKTSFFRCSLL 207
```

```
RESULT 12
RB8B_HUMAN
ID   RB8B_HUMAN     STANDARD;      PRT;   207 AA.
AC   Q92930; Q9P293;
DT   15-JUL-1999 (Rel. 38, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ras-related protein Rab-8B.
GN   RAB8B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RA   Seki N., Saito T.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 67-119 FROM N.A.
RC   TISSUE=Melanoma;
RX   MEDLINE=97182150; PubMed=9030196;
RA   Chen D., Guo J., Gahl W.A.;
RT   "RAB GTPases expressed in human melanoma cells.";
RL   Biochim. Biophys. Acta 1355:1-6(1997).
CC   -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC       NEUROTRANSMITTER RELEASE (BY SIMILARITY).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB038995; BAA92249.1; -.
DR   EMBL; BC020654; AAH20654.1; -.
DR   EMBL; U66624; AAC51199.1; -.
DR   HSSP; P05713; 3RAB.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein; Protein transport.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (BY SIMILARITY).
FT   LIPID       204    204       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   207 AA;  23584 MW;  5960993C0F87F944 CRC64;

  Query Match             43.6%;  Score 501;  DB 1;  Length 207;
  Best Local Similarity   49.2%;  Pred. No. 3.7e-37;
  Matches   98;  Conservative   40;  Mismatches   59;  Indels    2;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|| ::|||||
Db          4 TYDYLFKLLLIGDSGVGKTCLLFRFSEDAF-NTTFISTIGIDFKIRTIELDGKKIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||: |:  | |:|  ||  |:||||
```

```
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIKNWIRNIEEHASSDVERMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQA-DEPS 203
              ||: :| :  | || || :||: |||||||:  ||| ||  :|::: :   :  | |
Db        123 CDMNDKRQVSKERGEKLAIDYGIKFLETSAKSSANVEEAFFTLARDIMTKLNRKMNDSNS 182

Qy        204 FQIRDYVESQKKRSSCCSF 222
                  |:  : ||   ||
Db        183 AGAGGPVKITENRSKKTSF 201


RESULT 13
RIC1_ORYSA
ID   RIC1_ORYSA     STANDARD;      PRT;   202 AA.
AC   P40392;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ras-related protein RIC1.
GN   RIC1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Yamahoushi; TISSUE=Callus;
RX   MEDLINE=94009718; PubMed=8405471;
RA   Uchimiya H., Kidou S., Anai T., Umeda M., Aotsuka S., Tsuge T.,
RA   Kato A.;
RT   "Molecular structure of ras-related small GTP-binding protein genes
RT   of rice plants and GTPase activities of gene products in Escherichia
RT   coli.";
RL   FEBS Lett. 332:282-286(1993).
CC   -!- FUNCTION: POSSESS GTPASE ACTIVITY.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S66160; AAB28535.1; -.
DR   PIR; S38740; S38740.
DR   HSSP; P01112; 1PLK.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (PROBABLE).
FT   LIPID       200    200       PALMITATE (BY SIMILARITY).
FT   LIPID       201    201       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   202 AA;  22561 MW;  5A1CB5F93256B659 CRC64;

  Query Match             43.5%;  Score 500;  DB 1;  Length 202;
  Best Local Similarity   46.3%;  Pred. No. 4.4e-37;
  Matches   94;  Conservative   49;  Mismatches   54;  Indels    6;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
                 :| ||   |::|:|||:||||:| |::| | ::|  ::|:|:|:|||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLE-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQWLNEIDRYASENVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
              || |::  ||:  | |: || | |:||||||||   ||| ||: :| |:| |   |
Db        121 NKCDLAENRVVSYEAGKALADEIGIPFLETSAKDATNVEKAFMTMAGEIKNRMASQGRTN 180

Qy        199 ADEPSFQIRDYVESQKKRSSCCS 221
              | :|:  ::   :  ::|||||
Db        181 ASKPA-TVQMPRQPVAQQSSCCS 202


RESULT 14
SAS1_DICDI
ID   SAS1_DICDI     STANDARD;      PRT;   208 AA.
AC   P20790;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GTP-binding protein SAS1.
GN   SASA OR SAS1.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90220623; PubMed=2109188;
RA   Saxe S.A., Kimmel A.R.;
RT   "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
RT   with sequence similarities to essential genes in Saccharomyces
RT   cerevisiae.";
RL   Mol. Cell. Biol. 10:2367-2378(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89209367; PubMed=3149563;
RA   Saxe S.A., Kimmel A.R.;
RT   "Genes encoding novel GTP-binding proteins in Dictyostelium.";
RL   Dev. Genet. 9:259-265(1988).
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M34457; AAA33249.1; -.
DR   PIR; A34716; A34716.
DR   PIR; A61571; A61571.
DR   HSSP; P05713; 3RAB.
DR   DictyDb; DD05020; sasA.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Palmitate.
FT   NP_BIND      22     29       GTP (BY SIMILARITY).
FT   NP_BIND      70     74       GTP (BY SIMILARITY).
FT   NP_BIND     128    131       GTP (BY SIMILARITY).
FT   LIPID       207    207       PALMITATE (BY SIMILARITY).
FT   LIPID       208    208       PALMITATE (BY SIMILARITY).
SQ   SEQUENCE   208 AA;  23243 MW;  BF681EF5C810019E CRC64;

  Query Match             43.3%;  Score 498;  DB 1;  Length 208;
  Best Local Similarity   46.6%;  Pred. No. 6.8e-37;
```

```
  Matches   95;  Conservative   49;  Mismatches   50;  Indels   10;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              | :||   |::|:||:|||||:| |::| : :| : :|| |:||||: : : ::| |:|||
Db          9 SAAYDYLIKLLLIGDSGVGKSCLLLRFSEDSF-TPSFITTIGIDFKIRTIELEGKRIKLQ 67

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||::| |||| |  :||:||:|:: || ||| |:  | ::|   |  ||:|
Db         68 IWDTAGQERFRTITTAYYRGAMGILLVYDVTDEKSFGNIRNWIRNIEQHATDSVNKMLIG 127

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEP 202
              || ||: ::|: |  |::|| |||: |||||||  :||| ||:::||::| |     | |
Db        128 NKCDMAEKKVVDSSRGKSLADEYGIKFLETSAKNSINVEEAFISLAKDIKKR---MIDTP 184

Qy        203 SFQIR------DYVESQKKRSSCC 220
              : | :      :   :  |: :||
Db        185 NEQPQVVQPGTNLGANNNKKKACC 208


RESULT 15
YPT2_MAIZE
ID   YPT2_MAIZE     STANDARD;      PRT;   203 AA.
AC   Q05737;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   GTP-binding protein YPTM2.
GN   YPTM2.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Coleoptile;
RX   MEDLINE=92115746; PubMed=1731354;
RA   Palme K., Diefenthal T., Vingron M., Sander C., Schell J.;
RT   "Molecular cloning and structural analysis of genes from Zea mays
RT   (L.) coding for members of the ras-related ypt gene family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:787-791(1992).
CC   -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC       TRAFFIC (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: ITS EXPRESSION IS WEAK IN STEMS, HIGHER IN
CC       ROOTS, LEAVES AND COLEOPTILES, BUT HIGHEST IN FLOWERS.
CC   -!- PTM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
CC       BIOLOGICAL FUNCTION.
CC   -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X63278; CAA44919.1; -.
DR   PIR; B38202; B38202.
DR   HSSP; P05713; 3RAB.
DR   MaizeDB; 78605; -.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
FT   NP_BIND      15     22       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     121    124       GTP (BY SIMILARITY).
FT   DOMAIN       37     45       EFFECTOR REGION (PROBABLE).
FT   LIPID       200    200       PALMITATE (BY SIMILARITY).
FT   LIPID       201    201       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   203 AA;  22475 MW;  E241326E7ACD1B8A CRC64;

  Query Match             43.0%;  Score 495;  DB 1;  Length 203;
  Best Local Similarity   45.1%;  Pred. No. 1.2e-36;
  Matches   92;  Conservative   51;  Mismatches   53;  Indels    8;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              :| ||   |::|:||:|||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIIVYDVTDQESFNNVKQWLNEIDRYASDNVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
              ||:|::: :|: :|  :  | | |:||:|||||   ||: ||:|:|  :| |   |
Db        121 NKSDLTANKVVATETAKAFADEMGIPFMETSAKNATNVQQAFMAMAASIKDRMASQPAAA 180

Qy        199 -ADEPSFQIRDYVESQKKRSSCCS 221
               |   : |||    :||  :||||
Db        181 NARPATVQIRGQPVNQK--TSCCS 202


Search completed: June 18, 2003, 15:48:29
Job time : 33 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:37:02 ; Search time 71 Seconds
(without alignments)
418.520 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150	100.0	232	22	AAM41696	Human polypeptide
2	1144	99.5	226	22	AAU17119	Novel signal trans
3	1133	98.5	222	22	AAU17541	Novel signal trans
4	977	85.0	191	22	AAG67156	Amino acid sequenc
5	977	85.0	191	22	AAM39910	Human polypeptide
6	977	85.0	191	23	ABB76426	Human Ras protein
7	969	84.3	191	23	AAM52190	Human degranulatio
8	946	82.3	191	23	AAM52189	Mouse degranulatio
9	940	81.7	191	23	AAM52188	Degranulation regu
10	926	80.5	191	20	AAY30132	An exocytotic prot
11	770	67.0	158	22	ABG17244	Novel human diagno
12	697	60.6	666	22	ABB71578	Drosophila melanog
13	666	57.9	190	19	AAW80747	Human RAB protein,
14	535	46.5	139	22	AAU17123	Novel signal trans
15	514.5	44.7	221	23	ABP41333	Human ovarian anti
16	513.5	44.7	213	22	ABB11916	Human rab8 homolog
17	509	44.3	218	21	AAG19220	Arabidopsis thalia
18	508.5	44.2	204	22	ABB70670	Drosophila melanog
19	507	44.1	253	21	AAG53944	Arabidopsis thalia
20	507	44.1	254	21	AAG08687	Arabidopsis thalia
21	507	44.1	258	21	AAG30498	Arabidopsis thalia
22	505.5	44.0	215	21	AAG35215	Zea mays protein f
23	502	43.7	216	21	AAG08006	Arabidopsis thalia
24	501	43.6	203	21	AAG30499	Arabidopsis thalia
25	501	43.6	207	22	AAG67154	Amino acid sequenc
26	501	43.6	207	22	AAB92628	Human protein sequ
27	500	43.5	216	21	AAG08688	Arabidopsis thalia
28	500	43.5	216	21	AAG53945	Arabidopsis thalia
29	498	43.3	218	21	AAB56993	Human prostate can
30	498	43.3	218	22	ABG07266	Novel human diagno
31	497.5	43.3	202	21	AAG10858	Arabidopsis thalia
32	496.5	43.2	202	21	AAG07763	Arabidopsis thalia
33	494	43.0	200	21	AAB19165	Amino acid sequenc
34	494	43.0	200	21	AAB09979	Human Rab10 protei
35	494	43.0	200	22	AAB95340	Human protein sequ
36	488	42.4	224	21	AAG47826	Arabidopsis thalia
37	488	42.4	234	21	AAG47825	Arabidopsis thalia
38	487.5	42.4	201	21	AAB09982	Canine Rab10 prote
39	487.5	42.4	207	22	ABB71647	Drosophila melanog
40	487	42.3	209	22	ABG23365	Novel human diagno
41	483	42.0	218	21	AAG30710	Arabidopsis thalia
42	483	42.0	221	21	AAG30709	Arabidopsis thalia
43	483	42.0	246	21	AAB58196	Lung cancer associ
44	481.5	41.9	199	21	AAB09980	Human Rab10 protei
45	481.5	41.9	199	21	AAB09981	Human Rab10 protei

ALIGNMENTS

RESULT 1
AAM41696
ID AAM41696 standard; Protein; 232 AA.
XX
AC AAM41696;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6627.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI60852.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 6627; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 232 AA;

Query Match 100.0%; Score 1150; DB 22; Length 232;
Best Local Similarity 100.0%; Pred. No. 9.6e-120;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 69

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 129

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 189

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        190 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 232
```

RESULT 2
AAU17119
ID AAU17119 standard; Protein; 226 AA.
XX
AC AAU17119;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 684.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR N-PSDB; AAS27036.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 684; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 99.5%; Score 1144; DB 22; Length 226;
Best Local Similarity 99.6%; Pred. No. 4.3e-119;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 63

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ATVGIDFSNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 123

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 IRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNV 183

Qy        181 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        184 ELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 226
```

RESULT 3
AAU17541
ID AAU17541 standard; Protein; 222 AA.
XX
AC AAU17541;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 1106.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;

KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

```
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465460/50.
DR   N-PSDB; AAS27458.
XX
PT   Novel polypeptides useful for diagnosing, treating, preventing and/or
PT   prognosing disorders related to the proteins, including cancers, immune
PT   disorders and neuronal disorders  -
XX
PS   Claim 1; SEQ ID No 1106; 880pp; English.
XX
CC   The invention relates to novel isolated polypeptides (I), and
CC   polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC   diagnosing, preventing and treating diseases including immune system
CC   disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC   disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC   transplant rejections and graft versus host disease, infectious diseases
CC   (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC   other blood-related disorders (sickle cell anaemia), myeloproliferative
CC   disorders, primary haematopoietic disorders, hyperproliferative
CC   disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC   disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC   abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC   disorders (e.g. glomerulonephritis), cardiovascular disorders
CC   (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC   wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC   Addison's disease), reproductive system disorders, gastrointestinal
CC   disorder (inflammatory disorders), liver disorders (cirrhosis),
CC   as stimulators of B-cell responsiveness to pathogens, activators of
CC   T-cells, to induce higher affinity antibodies, and as a means to induce
CC   tumour proliferation in pathologies e.g. acquired immune deficiency
CC   syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC   pathway protein, amino acid sequences of the invention.
XX

  Query Match             98.5%;  Score 1133;  DB 22;  Length 222;
  Best Local Similarity   99.1%;  Pred. No. 7.1e-118;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIA 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIA 60

Qy         62 TVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNI 121
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TVGIDFXNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNI 120

Qy        122 RAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVE 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVE 180

Qy        182 LAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |||||||||||||||||||||||||||| ||||||||||||
Db        181 LAFLAIAKELKYRAGHQADEPSFQIRDYVXSQKKRSSCCSFM 222


RESULT 4
AAG67156
ID   AAG67156 standard; Protein; 191 AA.
XX
AC   AAG67156;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Amino acid sequence of human 32712 G-protein.
XX
KW   Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW   congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW   adult respiratory distress syndrome; Goodpasture's syndrome;
KW   chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
KW   liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
KW   cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
KW   brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
KW   acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
KW   chronic bacterial meningoencephalitis; multiple sclerosis;
KW   amyotropic lateral sclerosis; stroke; Huntington's disease.
XX
OS   Homo sapiens.
XX
PN   WO200164887-A2.
XX
PD   07-SEP-2001.
XX
PF   27-FEB-2001; 2001WO-US06292.
XX
PR   29-FEB-2000; 2000US-0185606.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers RA;
XX
DR   WPI; 2001-550182/61.
DR   N-PSDB; AAH75184.
XX
PT   Novel human small G-protein polypeptides and polynucleotides for
PT   treating lung disorders, liver disorders and brain disorders  -
XX
PS   Claim 8; Fig 26; 151pp; English.
XX
CC   The present sequence represents a human G-protein. The specification
CC   describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
CC   G-protein polypeptides and polynucleotides are useful as a target for
CC   diagnosis and treatment of  G-protein mediated or related disorders,
CC   and for identifying agonists and antagonists for diagnosis and
CC   treatment. They are useful for treating disorders of lung (e.g.
CC   congenital anomalies, pulmonary congestion, oedema, adult respiratory
CC   distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
CC   asthma, Goodpasture's syndrome and pulmonary hypertension), liver
CC   (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
CC   Wilson's disease, autoimmune hepatitis and hepatic failure), and
CC   brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
CC   meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
CC   bacterial meningoencephalitis, multiple sclerosis, amyotropic lateral
CC   sclerosis, stroke and Huntington's disease).
XX
SQ   Sequence   191 AA;

  Query Match             85.0%;  Score 977;  DB 22;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-100;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              |||||||||||
Db        181 QKKRSSCCSFM 191
```

```
RESULT 5
AAM39910
ID   AAM39910 standard; Protein; 191 AA.
XX
AC   AAM39910;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 3055.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI59066.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 3055; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   191 AA;

  Query Match             85.0%;  Score 977;  DB 22;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-100;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
               ||||||||||| 
Db        181 QKKRSSCCSFM 191


RESULT 6
ABB76426
ID   ABB76426 standard; Protein; 191 AA.
XX
AC   ABB76426;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Ras protein 3 (RASP-3).
XX
KW   Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;
KW   immunosuppressive; antiinflammatory; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1..19
FT                   /note= "transforming protein P21 family signature"
FT   Region          4..11
FT                   /note= "ATP/GTP-binding site, motif A (P-loop)"
FT   Region          22..38
FT                   /note= "transforming protein P21 family signature"
FT   Region          40..62
FT                   /note= "transforming protein P21 family signature"
FT   Region          63..80
FT                   /note= "Ran family signature sequence"
FT   Region          102..115
FT                   /note= "transforming protein P21 family signature"
FT   Region          130..168
FT                   /note= "Ran family signature sequence"
FT   Region          137..159
FT                   /note= "transforming protein P21 family signature"
XX
PN   US6391580-B1.
XX
PD   21-MAY-2002.
XX
PF   08-MAY-1998;   98US-0075454.
XX
PR   12-DEC-1996;   96US-0766551.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Tang YT,  Lal P,  Guegler KJ,  Corley NC,  Patterson C;
PI   Batra S,  Baughn MR;
XX
DR   WPI; 2002-498774/53.
DR   N-PSDB; ABN83696.
XX
PT   New isolated polynucleotides encoding Ras proteins designated RASP-1
PT   and RASP-4, for diagnosing, preventing and treating disorders
PT   associated with cell proliferation, particularly cancer and immune
PT   disorders  -
XX
PS   Example; Column 47-48; 34pp; English.
XX
CC   The present sequence is the protein sequence of novel human Ras
```

```
CC   protein 3 (RASP-3).  RASP-3 shares 72% sequence identity with
CC   Rab26.  It includes regions that resemble the signature sequences
CC   of the GTP-binding Ras superfamily, and family signatures of Ran,
CC   a subfamily of Ras proteins which function in nucleocytoplasmic
CC   transport, RNA synthesis, processing and export, and cell cycle
CC   checkpoint control, as well as family signatures of transforming
CC   protein P21, a subfamily of GTP-binding Ras proteins which have
CC   been implicated in a number of tumours.  Northern analysis showed
CC   expression of RASP-3 in haematopoietic and immunological cDNA
CC   libraries, all of which were associated with inflammation and the
CC   immune response.  The invention provides 7 novel human Ras proteins
CC   (RASP-1 to -7) and polynucleotides, expression vectors, host cells,
CC   antibodies, agonists and antagonists.  It also provides methods for
CC   diagnosing, treating or preventing disorders associated with RASP
CC   expression, especially cancer and immune disorders.
XX
SQ   Sequence   191 AA;

  Query Match             85.0%;  Score 977;  DB 23;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-100;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              ||||||||||| .
Db        181 QKKRSSCCSFM 191


RESULT 7
AAM52190
ID   AAM52190 standard; Protein; 191 AA.
XX
AC   AAM52190;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human degranulation regulator SEQ ID NO 3.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic.
XX
OS   Homo sapiens.
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
DR   N-PSDB; ABA02774.
XX
PT   Mast cell degranulation controller for treatment of allergies  -
XX
PS   Claim 11; Page 64-65; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence   191 AA;

  Query Match             84.3%;  Score 969;  DB 23;  Length 191;
  Best Local Similarity   99.0%;  Pred. No. 1.1e-99;
  Matches  189;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLGDSGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              |||||||||||:
Db        181 QKKRSSCCSFV 191


RESULT 8
AAM52189
ID   AAM52189 standard; Protein; 191 AA.
XX
AC   AAM52189;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Mouse degranulation regulator SEQ ID NO 2.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 83
FT                   /note= "Encoded by CAG"
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
DR   N-PSDB; ABA02773.
XX
PT   Mast cell degranulation controller for treatment of allergies  -
XX
PS   Claim 10; Page 63-64; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence   191 AA;

  Query Match             82.3%;  Score 946;  DB 23;  Length 191;
  Best Local Similarity   96.3%;  Pred. No. 4.2e-97;
```

```
  Matches  184;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              |||||:|||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 MLLGDSGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGARVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADVSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              |||||||||||||||:|||||||||||||||||||||||| | ||||||||||||
Db        121 IRSEDGETLAREYGVPFMETSAKTGMNVELAFLAIAKELKYRAGRQPDEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              ||||||||||:
Db        181 QKKRSSCCSFV 191


RESULT 9
AAM52188
ID   AAM52188 standard; protein; 191 AA.
XX
AC   AAM52188;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Degranulation regulator SEQ ID NO 1.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic.
XX
OS   Homo sapiens.
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 45
FT                   /label= Ala, Val
FT   Misc-difference 115
FT                   /label= Val, Met
FT   Misc-difference 138
FT                   /label= Met, Leu
FT   Misc-difference 165
FT                   /label= Arg, His
FT   Misc-difference 167
FT                   /label= Pro, Ala
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
XX
PT   Mast cell degranulation controller for treatment of allergies -
XX
PS   Claim 1; Page 63; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence   191 AA;

  Query Match             81.7%;  Score 940;  DB 23;  Length 191;
  Best Local Similarity   96.3%;  Pred. No. 2e-96;

  Matches  184;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              |||||:|||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 MLLGDSGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGXRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADXSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              |||||||||||||||| |||||||||||||||||||||||| | ||||||||||||
Db        121 IRSEDGETLAREYGVPFXETSAKTGMNVELAFLAIAKELKYRAGXQXDEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              ||||||||||:
Db        181 QKKRSSCCSFV 191


RESULT 10
AAY30132
ID   AAY30132 standard; Protein; 191 AA.
XX
AC   AAY30132;
XX
DT   27-OCT-1999  (first entry)
XX
DE   An exocytotic protein designated Exo2.
XX
KW   Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
KW   inflammation.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 5
FT                   /note= "encoded by GAN"
FT   Misc-difference 27
FT                   /note= "encoded by TTN"
XX
PN   WO9942586-A2.
XX
PD   26-AUG-1999.
XX
PF   23-FEB-1999;   99WO-US03909.
XX
PR   26-MAY-1998;   98US-0086650.
PR   23-FEB-1998;   98US-0075534.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Fisher J,  Huang B,  Lorens J,  Luo Y;
PI   Shen M;
XX
DR   WPI; 1999-518605/43.
DR   N-PSDB; AAX86720.
XX
PT   New exocytotic proteins useful for diagnosis and treatment of
PT   exocytosis-mediated conditions and in drug screening
XX
PS   Claim 12; Fig 7; 53pp; English.
XX
CC   The present sequence represents an exocytotic protein designated Exo2.
CC   The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC   secretory vesicles with the cellular plasma membrane) and can be
CC   administered therapeutically to treat or prevent exocytosis-mediated
CC   disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC   undesirable release of compounds via exocytosis e.g. inflammation
CC   mediated by the release compounds such as histamine. Exo2 is useful
CC   diagnostically and to produce antibodies useful to purify the proteins
CC   and therapeutically to reduce or eliminate the biological activity of
CC   the protein. Exo2 nucleic acids can be used therapeutically to increase
CC   Exo2 activity in cells by known gene therapy techniques. They can also
```

```
CC   be used to produce probes or primers to isolate Exo2 proteins from other
CC   organisms, especially humans. The nucleic acids, host cells and proteins
CC   are useful in screening assays to identify binding agents, especially
CC   drug screening assays to identify agonists and antagonists useful
CC   therapeutically to enhance or reduce Exo2 activity.
XX
SQ   Sequence   191 AA;

  Query Match             80.5%;  Score 926;  DB 20;  Length 191;
  Best Local Similarity   94.8%;  Pred. No. 7.2e-95;
  Matches  181;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              |||| :|||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db          1 MLLGXSGVGKTCFLIQFKDGAFLSGTXIATVGIDFRNKVVTVDGSRVKLQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              ||||||||||||||||||||:|||||||||||||||||||||||||||||:|||||
Db         61 RSVTHAYYRDAQALLLLYDITNQSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADVSSERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              |||||||||||||||||:|||||||||||||||||||||||| | ||||||||||||
Db        121 IRSEDGETLAREYGVPFMETSAKTGMNVELAFLAIAKELKYRAGRQPDEPSFQIRDYVES 180

Qy        213 QKKRSSCCSFM 223
              |||||||||:
Db        181 QKKRSSCCSFV 191


RESULT 11
ABG17244
ID   ABG17244 standard; Protein; 158 AA.
XX
AC   ABG17244;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #17235.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS81431.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 47603; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   158 AA;

  Query Match             67.0%;  Score 770;  DB 22;  Length 158;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-77;
  Matches  148;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI 69

Qy         61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDN 129

Qy        121 IRAWLTEIHEYAQRDVVIMLLGNKADMSS 149
              |||||||||||||:|||||||||||||||
Db        130 IRAWLTEIHEYAQKDVVIMLLGNKADMSS 158


RESULT 12
ABB71578
ID   ABB71578 standard; Protein; 666 AA.
XX
AC   ABB71578;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 41526.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL15681.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 41526; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
```

```
CC . capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   666 AA;

  Query Match             60.6%;  Score 697;  DB 22;  Length 666;
  Best Local Similarity   59.8%;  Pred. No. 1.6e-68;
  Matches  137;  Conservative   33;  Mismatches   49;  Indels   10;  Gaps    3;

Qy          3 GTPGAVAT---RDGEA------PERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGA 53
              |: || ||     | |         ::|     :|: |||::|||:|||||  ||:|:||
Db        433 GSEGASATLCKNAGRALIRMISSSKAPEEEDEFDIMGKVIMLGDSGVGKTSLLIRFRDGR 492

Qy         54 FLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDIT 113
              ::   |::|||||||||||| ||| |||||||||||||||||||||||||| |||||||:|
Db        493 YVPSYFLSTVGIDFRNKVVVVDGTRVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDVT 552

Qy        114 NKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMS-SERVIRSEDGETLAREYGVPFLET 172
              ||:::|||||| || |||| ||||:|:||||| | ||| :: |||| | ||: |||:||
Db        553 NKTTYDNIRAWLGEIREYAQEDVVIVLIGNKADCSGSERQVKREDGERLGREHNVPFMET 612

Qy        173 SAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSSCCS 221
              |||||:||||:| |:|::|| |     |:  | : |:|     |  | |:
Db        613 SAKTGLNVELSFTAVARQLKSRGYEHGDDGKFNVHDFVRDNTKARSVCA 661


RESULT 13
AAW80747
ID   AAW80747 standard; Protein; 190 AA.
XX
AC   AAW80747;
XX
DT   17-DEC-1998  (first entry)
XX
DE   Human RAB protein, SRAB.
XX
KW   Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW   Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW   viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW   diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW   hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;
KW   polycystic renal disease.
XX
OS   Homo sapiens.
XX
PN   WO9842839-A1.
XX
PD   01-OCT-1998.
XX
PF   25-MAR-1998;   98WO-US05871.
XX
PR   26-MAR-1997;   97US-0824873.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Guegler KJ,  Hillman JL;
XX
DR   WPI; 1998-542282/46.
DR   N-PSDB; AAV65197.
XX
PT   New human Rab protein, SRAB - useful for treating disorders
PT   associated with SRAB expression, including vesicle trafficking,
PT   viral infection, and cancer
XX
PS   Claim 1; Fig 1; 31pp; English.
XX
CC   This sequence is the human RAB protein, SRAB, of the invention. The
CC   SRAB protein can be used in compositions useful in the diagnosis,
CC   prevention, or treatment of disorders associated with vesicle trafficking
CC   (including epilepsy, Huntington's disease, Parkinson's disease and
CC   schizophrenia), cancer, or viral infection. Host cells containing the DNA
CC   sequence can be used to produce SRAB recombinantly. The viral infections
CC   that can be treated using SRAB include those caused by retroviruses
CC   particularly HIV and HTLV, hepadnaviruses, particularly hepatitis C,
CC   hantaviruses, herpesviruses, and arboviruses. The cancers that can be
CC   treated include adenocarcinoma, leukaemia, lymphoma, melanoma and
CC   sarcoma, particularly cancers of the endocrine, gastrointestinal and
CC   nervous systems and cancers of the adrenal gland, brain, breast, colon,
CC   oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,
CC   prostate, salivary gland, stomach, thyroid, and uterus. The protein can
CC   also be used to treat conditions such as autoimmune sialosis,
CC   choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,
CC   hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme
CC   deficiency, pancreatitis, and polycystic renal disease.
XX
SQ   Sequence   190 AA;

  Query Match             57.9%;  Score 666;  DB 19;  Length 190;
  Best Local Similarity   67.0%;  Pred. No. 7.4e-66;
  Matches  126;  Conservative   25;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:||:|||||| |   :   |  |       |||||||: ||||:||||:|||||||||
Db          1 MLVGDSGVGKTCLLGAIQGWCFPGGDLHLHRSIDFRNKVLDVDGVKVKLQMWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||| |||||||:|||:|||||:|||||||||| || :|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGNKVDSAHERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: |||| ||:|||:||:||||||||:||:||| ||||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : : :|||
Db        181 EGRGASCC 188


RESULT 14
AAU17123
ID   AAU17123 standard; Protein; 139 AA.
XX
AC   AAU17123;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Novel signal transduction pathway protein, Seq ID 688.
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW   antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis C; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01312.
XX
PR   31-JAN-2000; 2000US-0179065.
```

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR N-PSDB; AAS27040.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -

```
XX
PS   Claim 1; SEQ ID No 688; 880pp; English.
XX
CC   The invention relates to novel isolated polypeptides (I), and
CC   polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC   diagnosing, preventing and treating diseases including immune system
CC   disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC   disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC   transplant rejections and graft versus host disease, infectious diseases
CC   (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC   other blood-related disorders (sickle cell anaemia), myeloproliferative
CC   disorders, primary haematopoietic disorders, hyperproliferative
CC   disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC   disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC   abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC   disorders (e.g. glomerulonephritis), cardiovascular disorders
CC   (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC   wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC   Addison's disease), reproductive system disorders, gastrointestinal
CC   disorder (inflammatory disorders), liver disorders (cirrhosis),
CC   as stimulators of B-cell responsiveness to pathogens, activators of
CC   T-cells, to induce higher affinity antibodies, and as a means to induce
CC   tumour proliferation in pathologies e.g. acquired immune deficiency
CC   syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC   pathway protein, amino acid sequences of the invention.
XX

  Query Match             46.5%;  Score 535;  DB 22;  Length 139;
  Best Local Similarity   71.5%;  Pred. No. 1.9e-51;
  Matches   98;  Conservative   19;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          84 WDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGN 143
               ||||||||||||||||||| |||||||:|||:|||||:|||||||||||| || :|||||
Db           1 WDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQHDVALMLLGN 60

Qy         144 KADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPS 203
               | | : |||:: |||| ||:|||:||:|||||||:||:||| ||||||| |:     ||
Db          61 KVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELKQRSMKAPSEPR 120

Qy         204 FQIRDYVESQKKRSSCC 220
               |:: |||: : : :|||
Db         121 FRLHDYVKREGRGASCC 137


RESULT 15
ABP41333
ID   ABP41333 standard; Protein; 221 AA.
XX
AC   ABP41333;
XX
DT   23-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HCGMA67, SEQ ID NO:2465.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ54410.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 11; SEQ ID No 2465; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents a human ovarian antigen of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   221 AA;

  Query Match             44.7%;  Score 514.5;  DB 23;  Length 221;
  Best Local Similarity   48.3%;  Pred. No. 7.4e-49;
  Matches  100;  Conservative   43;  Mismatches   59;  Indels    5;  Gaps    2;

Qy          16 PERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVD 75
               |  :    : :||   |::|:||:|||||| | :| : || : |||:|:||||: : : :|
Db           9 PGSTHASAKTYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELD 67

Qy          76 GVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRD 135
               | |:|||||||||||||::| |||| |  ::|:|||||: ||||||  |:  | |:|  |
Db          68 GKRIKLQIWDTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASAD 127

Qy         136 VVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
               |  |:|||| |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :
Db         128 VEKMILGNKCDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKM 187

Qy         196 GHQAD----EPSFQIRDYVESQKKRSS 218
                 : :    : | |        |:|||
Db         188 DKKLEGNSPQGSNQGVKITPDQQKRSS 214


Search completed: June 18, 2003, 15:47:55
Job time : 77 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:44:07 ; Search time 40 Seconds
(without alignments)
535.949 Million cell updates/sec

Title: US-09-817-199B-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSSCCSFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	742	64.5	190	2	JC2528	GTP-binding protei
2	533	46.3	196	2	T15123	hypothetical prote
3	517.5	45.0	203	2	B34716	GTP-binding protei
4	516	44.9	206	2	I78851	GTP-binding protei
5	513.5	44.7	207	2	B49647	GTP-binding protei
6	513.5	44.7	207	2	B36364	GTP-binding protei
7	512.5	44.6	215	2	T14565	GTP-binding protei
8	507	44.1	216	2	T48378	GTP-binding protei
9	507	44.1	258	2	B86153	ARA-5 [imported] -
10	506	44.0	216	2	S33900	GTP-binding protei
11	505.5	44.0	201	2	T28971	hypothetical prote
12	500	43.5	202	2	S38740	GTP-binding protei
13	500	43.5	216	2	T45901	GTPase AtRAB8 - Ar
14	498	43.3	208	2	A34716	GTP-binding protei
15	497.5	43.3	202	2	S41430	GTP-binding protei
16	497.5	43.3	209	2	B38625	GTP-binding protei
17	497.5	43.3	215	2	S57478	GTP-binding protei
18	495	43.0	203	2	B38202	GTP-binding protei
19	495	43.0	216	2	S57471	GTP-binding protei
20	495	43.0	216	2	JS0640	GTP-binding protei
21	494	43.0	200	2	D36364	GTP-binding protei
22	490.5	42.7	200	2	A38625	GTP-binding protei
23	490.5	42.7	215	2	S57462	GTP-binding protei
24	489	42.5	222	2	T14405	small GTP-binding
25	488	42.4	203	2	S34253	GTP-binding protei
26	487	42.3	200	2	S12790	GTP-binding protei
27	486	42.3	200	2	B42148	GTP-binding protei
28	486	42.3	202	2	S72515	GTP-binding protei
29	486	42.3	203	2	JC4105	GTP-binding protei
30	485	42.2	196	2	PS0279	GTP-binding protei
31	484.5	42.1	215	2	S57474	GTP-binding protei
32	483	42.0	221	2	H71444	GTP-binding protei
33	481.5	41.9	224	2	T33855	hypothetical prote
34	479.5	41.7	204	2	JC7589	Sec4p homolog - ye
35	479.5	41.7	208	2	A38202	GTP-binding protei
36	478.5	41.6	206	2	T14391	GTP-binding protei
37	478	41.6	201	2	S06147	GTP-binding protei
38	478	41.6	203	2	JC1247	GTP-binding protei
39	477.5	41.5	205	2	S38339	GTP-binding protei
40	476.5	41.4	205	2	T33781	hypothetical prote
41	476	41.4	217	2	S36365	GTP-binding protei
42	473.5	41.2	203	2	A49647	GTP-binding protei
43	468.5	40.7	201	2	S39565	GTP-binding protei
44	463.5	40.3	203	2	S51495	GTP-binding protei
45	459.5	40.0	203	2	S30096	GTP-binding protei

ALIGNMENTS

RESULT 1
JC2528
GTP-binding protein Rab26 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001
C;Accession: JC2528
R;Wagner, A.C.C.; Strowski, M.Z.; Goeke, B.; Williams, J.A.
Biochem. Biophys. Res. Commun. 207, 950-956, 1995
A;Title: Molecular cloning of a new member of the Rab protein family, Rab26, from rat
A;Reference number: JC2528; MUID:95169156; PMID:7864900
A;Accession: JC2528
A;Molecule type: mRNA
A;Residues: 1-190 <WAG>
A;Cross-references: GB:T08521
A;Note: The authors translated the codon GTG for residue 49 as Leu, GAA for residue 1
C;Genetics:
A;Gene: Rab26
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;4-11/Region: nucleotide-binding motif A (P-loop)
F;52-58/Region: GTP binding #status predicted
F;111-114/Region: GTP binding #status predicted
F;139-143/Region: GTP binding #status predicted

Query Match 64.5%; Score 742; DB 2; Length 190;
Best Local Similarity 72.9%; Pred. No. 3.4e-57;
Matches 137; Conservative 28; Mismatches 23; Indels 0; Gaps 0;

```
Qy         33 MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERF 92
              ||:|!:|||||| |::|||||||:||||:||||||||||: |||::||:||||||||||
Db          1 MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDGMKVKVQIWDTAGQERF 60

Qy         93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERV 152
              |||||||||||| |||||||||||| |||||:|||||| ||||:|||:|||||| | : |||
Db         61 RSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQDVVLMLLGNKVDSTQERV 120

Qy        153 IRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
              :: ||| ||:|||:||:|||||:|:||:||| ||||||| |:     || |:: |||:
Db        121 VKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELKQRSTKAPSEPRFRLHDYVKR 180

Qy        213 QKKRSSCC 220
              : :  |||
Db        181 EGRGVSCC 188
```

RESULT 2
T15123
hypothetical protein W01H2.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T15123
R;Minx, P.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid W01H2.
A;Reference number: Z18296
A;Accession: T15123
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-196 <MIN>
A;Cross-references: EMBL:AF000192; NID:g1946982; PID:g1946985; PIDN:AAB52888.1; GSPDB:GN
A;Experimental source: strain Bristol N2; clone W01H2
C;Genetics:
A;Gene: CESP:W01H2.3
A;Map position: X
A;Introns: 41/3; 55/3; 95/3; 134/3; 160/2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
F;3-119/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 533; DB 2; Length 196;
Best Local Similarity 54.5%; Pred. No. 5.2e-39;
Matches 109; Conservative 34; Mismatches 41; Indels 16; Gaps 5;

```
Qy         31 KVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQE 90
              |||||||:  |||| ||::|||||||:  ||:|||||:|||::|:   :||||||||||||
Db          4 KVMLLGDSCTGKTCLLIRYKDGAFLNNNFISTVGIDYRNKLITMGDKKVKLQIWDTAGQE 63

Qy         91 RFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSE 150
              |||||| :||||| ||||:||| |::||:| | ||::| || :  | : |:||| |:
Db         64 RFRSVTTSYYRDADALLLVYDIANRASFENCRNWLSQIKEYGKEAVQVTLVGNKCDL--P 121

Qy        151 RVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAK---ELKY---RAGHQADEPSF 204
              | : :::|: ||  | :||:||||||| ||: ||| :|:   :|||     | ||  |
Db        122 RAVPTDEGKRLAEAYQIPFMETSAKTGFNVDRAFLGLAERMLKLKYGFVPGGEMADTIS- 180

Qy        205 QIRDYVESQKKR--SSCCSF 222
                    |  ||   : ||:|
Db        181 -----VADTKKPEIARCCTF 195
```

RESULT 3
B34716
GTP-binding protein SAS2 - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 02-Feb-2001
C;Accession: B34716; B61571
R;Saxe, S.A.; Kimmel, A.R.
Mol. Cell. Biol. 10, 2367-2378, 1990
A;Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum with seque
A;Reference number: A34716; MUID:90220623; PMID:2109188
A;Accession: B34716
A;Molecule type: DNA
A;Residues: 1-203 <SAX>
A;Cross-references: GB:M34457
R;Saxe, S.A.; Kimmel, A.R.
Dev. Genet. 9, 259-265, 1988
A;Title: Genes encoding novel GTP-binding proteins in Dictyostelium.
A;Reference number: A61571; MUID:89209367; PMID:3149563
A;Accession: B61571
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-203 <SA2>
C;Genetics:
A;Gene: SAS2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;16-131/Domain: translation elongation factor Tu homology <ETU>
F;22-29/Region: nucleotide-binding motif A (P-loop)
F;128-131/Region: GTP-binding NKXD motif
F;158-160/Region: GTP-binding SAK/L motif
F;202,203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 45.0%; Score 517.5; DB 2; Length 203;
Best Local Similarity 47.1%; Pred. No. 1.2e-37;
Matches 97; Conservative 50; Mismatches 54; Indels 5; Gaps 3;

```
Qy         15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTV 74
              :|  : |   :||   |::|:||:||||:| |::| ||:| : :||||:||||: : : :
Db          3 SPATNKPA--AYDFLVKLLLIGDSGVGKSCLLLRFSDGSF-TPSFIATIGIDFKIRTIEL 59

Qy         75 DGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR 134
              :| |:|||||||||||||::| |||| |  :||:||:|:: || :|| |:  | ::|
Db         60 EGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDEKSFGSIRNWIRNIEQHASD 119

Qy        135 DVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR 194
               |  ||:||| ||: ::|: |  |::|| ||||: ||||||  :||| ||: :||::| |
Db        120 SVNKMLIGNKCDMTEKKVVDSSRGKSLADEYGIKFLETSAKNSVNVEEAFIGLAKDIKKR 179

Qy        195 AGHQADEPSFQIRDYVESQKKRSSCC 220
                    ::|   |   :    |:::||
Db        180 MIDTPNDPDHTI--CITPNNKKNTCC 203
```

RESULT 4
I78851
GTP-binding protein MEL - mouse
N;Alternate names: gene MEL protein
C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Feb-2001
C;Accession: I78851
R;Nimmo, E.R.; Sanders, P.G.; Padua, R.A.; Hughes, D.; Williamson, R.; Johnson, K.J.
Oncogene 6, 1347-1351, 1991
A;Title: The MEL gene: a new member of the RAB/YPT class of RAS-related genes.
A;Reference number: I58355; MUID:91360267; PMID:1886711
A;Accession: I78851
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-206 <RES>
A;Cross-references: GB:S53270; NID:g234747; PIDN:AAB19682.1; PID:g234748
C;Genetics:
A;Gene: MEL
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.9%; Score 516; DB 2; Length 206;
Best Local Similarity 49.5%; Pred. No. 1.7e-37;
Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;

```
Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:|||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ----AD 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :        |
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKNWKATAA 182

Qy        201 EPSFQIRDYVESQKKRSSC-CSFM 223
                 |  ::  || ||: |   || :
Db        183 GSSHGVKITVEQQKRTSFFRCSLL 206
```

RESULT 5
B49647
GTP-binding protein rab8 - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Jan-2001
C;Accession: B49647; S36817
R;Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Lou
J. Cell Biol. 124, 101-115, 1994
A;Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c

```
A;Reference number: A49647; MUID:94124602; PMID:8294494
A;Accession: B49647
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-207 <ZAH>
A;Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318
R;Joberty, G.; Tavitian, A.; Zahraoui, A.
FEBS Lett. 330, 323-328, 1993
A;Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.
A;Reference number: S36817; MUID:93387463; PMID:8375503
A;Accession: S36817
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 175-186 <JOB>
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

  Query Match          44.7%;  Score 513.5;  DB 2;  Length 207;
  Best Local Similarity  50.0%;  Pred. No. 2.8e-37;
  Matches   99;  Conservative   41;  Mismatches   53;  Indels    5;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:|||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        183 QGSNQGVKITPDQQKRSS 200


RESULT 6
B36364
GTP-binding protein rab8 - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2001
C;Accession: B36364; S15604
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.
A;Reference number: A36364; MUID:91061765; PMID:2123294
A;Accession: B36364
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-207 <CHA>
A;Cross-references: GB:X56385; NID:g920; PIDN:CAB56776.1; PID:g6006436
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

  Query Match          44.7%;  Score 513.5;  DB 2;  Length 207;
  Best Local Similarity  50.0%;  Pred. No. 2.8e-37;
  Matches   99;  Conservative   41;  Mismatches   53;  Indels    5;  Gaps    2;

Qy         25 SYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIW 84
              :||   |::|:||:|||||| | :| : || : |||:|:|||||: : : :|| |:|||||
Db          4 TYDYLFKLLLIGDSGVGKTCVLFRFSEDAF-NSTFISTIGIDFKIRTIELDGKRIKLQIW 62

Qy         85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144
              |||||||||::| |||| |  ::|:|||||: |||||| |:  | |:|  ||  |:||||
Db         63 DTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIRNWIRNIEEHASADVEKMILGNK 122

Qy        145 ADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQAD---- 200
               |:: :| :  | || || :||: |:|||||  :||| ||  :|:::| :   : :
Db        123 CDVNDKRQVSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEGNSP 182

Qy        201 EPSFQIRDYVESQKKRSS 218
              : | |       |:||||
Db        183 QGSNQGVKITPDQQKRSS 200


RESULT 7
T14565
GTP-binding protein - beet
N;Alternate names: small G protein
C;Species: Beta vulgaris (beet)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C;Accession: T14565
R;Dallery, E.; Oulef, S.; Ben Jilany, K.E.; Kerckaert, J.; Hagege, D.
submitted to the EMBL Data Library, May 1995
A;Description: Molecular cloning and structural analysis of cDNAs that encode three s
A;Reference number: Z18142
A;Accession: T14565
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-215 <DAL>
A;Cross-references: EMBL:Z49152; NID:g974775; PID:g974776
A;Experimental source: strain D100 KS 38080
C;Function:
A;Description: GTP-binding
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
F;16-131/Domain: translation elongation factor Tu homology <ETU>

  Query Match          44.6%;  Score 512.5;  DB 2;  Length 215;
  Best Local Similarity  47.6%;  Pred. No. 3.5e-37;
  Matches  101;  Conservative   47;  Mismatches   53;  Indels   11;  Gaps    5;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |::|:||:||||:| |::| ||:| : :|| |:|||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
               :|:|||||  |:| : :  |: || |||: | |||||| :|||  | :||:::| |
Db        122 NKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVEEVFFSIARDIKQRL 181

Qy        196 G----HQADEPSFQIRDYVESQKK---RSSCC 220
                   |  :||  |:   :|  :   :|:||
Db        182 ADSDTRQEAQPSITIKPADQSGNQAAAKSACC 213


RESULT 8
T48378
GTP-binding protein-like - Arabidopsis thaliana
N;Alternate names: protein F12E4.300
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
C;Accession: T48378
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke,
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24492
A;Accession: T48378
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-216 <BEV>
A;Cross-references: EMBL:AL162751
A;Experimental source: cultivar Columbia; BAC clone F12E4
```

```
C;Genetics:
A;Map position: 5
A;Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3
A;Note: F12E4.300
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
F;16-131/Domain: translation elongation factor Tu homology <ETU>

  Query Match             44.1%;  Score 507;  DB 2;  Length 216;
  Best Local Similarity   47.5%;  Pred. No. 1.1e-36;
  Matches  103;  Conservative   46;  Mismatches   56;  Indels   12;  Gaps    4;

Qy         15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTV 74
              || |:      ||   |:::|:|||:||||:| |::| |  | : :|| |:||||: : | :
Db          4 APARA---RSDYDYLIKLLLIGDSGVGKSCLLLRFSDDTFTT-SFITTIGIDFKIRTVEL 59

Qy         75 DGVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR 134
              || |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|
Db         60 DGKRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWMKNIEQHASD 119

Qy        135 DVVIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKY 193
              :|  :|:|||||  |:| : :  |: || |||: | ||||||  :|||  |::|||::|
Db        120 NVNKILVGNKADMDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVENVFMSIAKDIKQ 179

Qy        194 R-------AGHQADEPSFQIRDYVESQKKRSSCCSFM 223
              |       |  |  : : |      |  ::|:|||::
Db        180 RLTETDTKAEPQGIKITKQDTAASSSTAEKSACCSYV 216


RESULT 9
B86153
ARA-5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Aug-2002
C;Accession: B86153
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86153
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <STO>
A;Cross-references: GB:AE005172; NID:g2317906; PIDN:AAC24370.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

  Query Match             44.1%;  Score 507;  DB 2;  Length 258;
  Best Local Similarity   45.5%;  Pred. No. 1.4e-36;
  Matches   95;  Conservative   50;  Mismatches   56;  Indels    8;  Gaps    3;

Qy         18 RSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGV 77
              ||   :| ||   |:::|:||:||||:| |::| | :::  ::|:|:|:||: : |  ||
Db         52 RSKTMNPEYDYLFKLLLIGDSGVGKSCLLLRFSDDSYVE-SYISTIGVDFKIRTVEQDGK 110

Qy         78 RVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVV 137
              :|||||||||||||::| :||| |  ::::||:|::  ||:|::  ||:||  ||  :|
Db        111 TIKLQIWDTAGQERFRTITSSYYRGAHGIIIVYDVTDEESFNNVKQWLSEIDRYASDNVN 170

Qy        138 IMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR--- 194
              :|:|||:|::  | |  |  :  | | |:||:|||||   ||| ||:|::  :| |
Db        171 KLLVGNKSDLTENRAIPYETAKAFADEIGIPFMETSAKDATNVEQAFMAMSASIKERMAS 230

Qy        195 --AGHQADEPSFQIRDYVESQKKRSSCCS 221
                ||: |  |: |||    :||  : |||
Db        231 QPAGNNARPPTVQIRGQPVAQK--NGCCS 257


RESULT 10
S33900
GTP-binding protein ypt2 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001
C;Accession: S33900; JQ2233
R;Fleming, A.J.; Mandel, T.; Roth, I.; Kuhlemeier, C.
Plant Cell 5, 297-309, 1993
A;Title: The patterns of gene expression in the tomato shoot apical meristem.
A;Reference number: S33899; MUID:93222691; PMID:8467223
A;Accession: S33900
A;Molecule type: mRNA
A;Residues: 1-216 <FL2>
A;Cross-references: GB:X69980; NID:g313028; PIDN:CAA49600.1; PID:g313029
C;Genetics:
A;Gene: ypt2
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
F;16-131/Domain: translation elongation factor Tu homology <ETU>
F;22-29/Region: nucleotide-binding motif A (P-loop)
F;128-131/Region: GTP-binding NKXD motif
F;159-161/Region: GTP-binding SAK/L motif

  Query Match             44.0%;  Score 506;  DB 2;  Length 216;
  Best Local Similarity   46.9%;  Pred. No. 1.3e-36;
  Matches  100;  Conservative   47;  Mismatches   54;  Indels   12;  Gaps    5;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||      ||   |:::|:|||||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDTGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
               |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|  :|
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDNV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYR- 194
               :|:||||||  |:| : :  |: || |||: | |||||| :|||  | :| |::| |
Db        122 NKILVGNKADMDESKRAVPTSKGQALADEYGIKFFETSAKTNLNVEEVFFSIGKDIKQRL 181

Qy        195 --AGHQADEPSFQIRDYVES-----QKKRSSCC 220
                :  : :  | :|    ::         ::||||
Db        182 SESDSKTEPQSIRINQSDQAGTAGQGAQKSSCC 214


RESULT 11
T28971
hypothetical protein T23H2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C;Accession: T28971
R;Wamsley, P.; Bradshaw, H.
submitted to the EMBL Data Library, January 1997
A;Description: The sequence of C. elegans cosmid T23H2.
A;Reference number: Z20549
A;Accession: T28971
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-201 <WAM>
A;Cross-references: EMBL:U80033; PIDN:AAC48200.1; GSPDB:GN00019; CESP:T23H2.5
A;Experimental source: strain Bristol N2; clone T23H2
C;Genetics:
A;Gene: CESP:T23H2.5
A;Map position: 1
A;Introns: 43/1; 147/2; 173/3
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

  Query Match             44.0%;  Score 505.5;  DB 2;  Length 201;
  Best Local Similarity   49.7%;  Pred. No. 1.3e-36;
  Matches   98;  Conservative   39;  Mismatches   57;  Indels    3;  Gaps    2;

Qy         26 YDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWD 85
```

```
              ||:  |::|:||:|||||| | :| | || : |||:|:||||: | : : | ::||||||
Db          6 YDMLFKLLLIGDSGVGKTCILYRFSDDAF-NTTFISTIGIDFKIKTIELKGKKIKLQIWD 64

Qy         86 TAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKA 145
              ||||||| ::| :||| |  ::|:|||||  |||||  ||  | |:|  ||| |:||||
Db         65 TAGQERFHTITTSYYRGAMGIMLVYDITNAKSFDNIAKWLRNIDEHASEDVVKMILGNKC 124

Qy        146 DMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQ 205
              |||  ||:  | || :|:::|: | |||||  ::|: ||  :|: :  :     || |
Db        125 DMSDRRVVSRERGEKIAQDHGISFHETSAKLNVHVDTAFYDLAEAILAKMPDSTDEQSRD 184

Qy        206 IRDYVESQKKRSS--CC 220
                : |: |:: ||  ||
Db        185 TVNPVQPQRQSSSGGCC 201


RESULT 12
S38740
GTP-binding protein - rice
N;Alternate names: ras-related small GTP-binding protein
C;Species: Oryza sativa (rice)
C;Date: 27-May-1994 #sequence_revision 27-Feb-1997 #text_change 02-Feb-2001
C;Accession: S38740
R;Kidou, S.; Anai, T.; Umeda, M.; Aotsuka, S.; Tsuge, T.; Kato, A.; Uchimiya, H.
FEBS Lett. 332, 282-286, 1993
A;Title: Molecular structure of ras-related small GTP-binding protein genes of rice plan
A;Reference number: S38740; MUID:94009718; PMID:8405471
A;Accession: S38740
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-202 <KID>
A;Cross-references: GB:S66160; NID:g432606; PIDN:AAB28535.1; PID:g432607
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             43.5%;  Score 500;  DB 2;  Length 202;
  Best Local Similarity   46.3%;  Pred. No. 4e-36;
  Matches   94;  Conservative   49;  Mismatches   54;  Indels    6;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              :| ||   |::|:||:||||:| |::| | ::|  ::|:|:|:||: : |  ||  :|||
Db          2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYLE-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
              |||||||||||::| :||| |  ::::||:|:: ||:|:: || ||  ||  :|  :|:|
Db         61 IWDTAGQERFRTITSSYYRGAHGIIVVYDVTDQESFNNVKQWLNEIDRYASENVNKLLVG 120

Qy        143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
              || |::  ||:  | |: || | |:||||||||   ||| ||: :| |:| |   |
Db        121 NKCDLAENRVVSYEAGKALADEIGIPFLETSAKDATNVEKAFMTMAGEIKNRMASQGRTN 180

Qy        199 ADEPSFQIRDYVESQKKRSSCCS 221
              | :|:  ::   :   ::||||
Db        181 ASKPA-TVQMPRQPVAQQSSCCS 202


RESULT 13
T45901
GTPase AtRAB8 - Arabidopsis thaliana
N;Alternate names: protein F4P12.310
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C;Accession: T45901
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23016
A;Accession: T45901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-216 <BLO>
A;Cross-references: EMBL:AL132966
A;Experimental source: cultivar Columbia; BAC clone F4P12
C;Genetics:
A;Map position: 3
A;Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3
A;Note: F4P12.310
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
F;16-131/Domain: translation elongation factor Tu homology <ETU>

  Query Match             43.5%;  Score 500;  DB 2;  Length 216;
  Best Local Similarity   46.5%;  Pred. No. 4.3e-36;
  Matches   99;  Conservative   45;  Mismatches   57;  Indels   12;  Gaps    5;

Qy         19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG 76
              :||     ||   |::|:||:||||:| |::| ||:| : :|| |:||||: : : :||
Db          3 APPARARADYDYLIKLLLIGDSGVGKSCLLLRFSDGSFTT-SFITTIGIDFKIRTIELDG 61

Qy         77 VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136
              |:|||||||||||||::| |||| |  :||:||:|::|||:||| |:  | ::|   |
Db         62 KRIKLQIWDTAGQERFRTITTAYYRGAMGILLVYDVTDESSFNNIRNWIRNIEQHASDSV 121

Qy        137 VIMLLGNKADM-SSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRA 195
               :|:|||||  |:| :    |: || |||: | |||||| :|||  | :|||::| |
Db        122 NKILVGNKADMDESKRAVPKSKGQALADEYGMKFFETSAKTNLNVEEVFFSIAKDIKQRL 181

Qy        196 GH---QADEPSFQIRDY-----VESQKKRSSCC 220
                 :|:  : :|          ::|:||
Db        182 ADTDARAEPQTIKINQSDQGAGTSQATQKSACC 214


RESULT 14
A34716
GTP-binding protein SAS1 - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 02-Feb-2001
C;Accession: A34716; A61571
R;Saxe, S.A.; Kimmel, A.R.
Mol. Cell. Biol. 10, 2367-2378, 1990
A;Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum with se
A;Reference number: A34716; MUID:90220623; PMID:2109188
A;Accession: A34716
A;Molecule type: mRNA
A;Residues: 1-208 <SAX>
A;Cross-references: GB:M34456
R;Saxe, S.A.; Kimmel, A.R.
Dev. Genet. 9, 259-265, 1988
A;Title: Genes encoding novel GTP-binding proteins in Dictyostelium.
A;Reference number: A61571; MUID:89209367; PMID:3149563
A;Accession: A61571
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-208 <SA2>
C;Genetics:
A;Gene: SAS1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loo
F;16-131/Domain: translation elongation factor Tu homology <ETU>
F;22-29/Region: nucleotide-binding motif A (P-loop)
F;128-131/Region: GTP-binding NKXD motif
F;158-160/Region: GTP-binding SAK/L motif
F;207,208/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

  Query Match             43.3%;  Score 498;  DB 2;  Length 208;
  Best Local Similarity   46.6%;  Pred. No. 6.2e-36;
  Matches   95;  Conservative   49;  Mismatches   50;  Indels   10;  Gaps    3;

Qy         23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
              | :|:|   |::|:||:||||:| |::| : :| : :|| |:||||: : : ::| |:|||
Db          9 SAAYDYLIKLLLIGDSGVGKSCLLLRFSEDSF-TPSFITTIGIDFKIRTIELEGKRIKLQ 67

Qy         83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
```

```
                ||||||||||||::| |||| |  :||:||:|:: || ||| |:  | ::|   |  ||:|
Db          68 IWDTAGQERFRTITTAYYRGAMGILLVYDVTDEKSFGNIRNWIRNIEQHATDSVNKMLIG 127

Qy         143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEP 202
                || ||: ::|: |  |::|| ||||: ||||||| :||| ||:::||::| |     | |
Db         128 NKCDMAEKKVVDSSRGKSLADEYGIKFLETSAKNSINVEEAFISLAKDIKKR---MIDTP 184

Qy         203 SFQIR------DYVESQKKRSSCC 220
                : | :      :   :  |: :||
Db         185 NEQPQVVQPGTNLGANNNKKKACC 208


RESULT 15
S41430
GTP-binding protein, ras-like (clone vfa-ypt1) - fava bean
N;Alternate names: guanine nucleotide regulatory protein
C;Species: Vicia faba (fava bean)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Feb-2001
C;Accession: S41430
R;Saalbach, G.; Thielmann, J.
submitted to the EMBL Data Library, January 1994
A;Description: Sequences of cDNA clones from cotyledons of Vicia faba encoding ypt/rab-r
A;Reference number: S41430
A;Accession: S41430
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-202 <SAA>
A;Cross-references: EMBL:Z29590; NID:g452358; PIDN:CAA82707.1; PID:g452359
A;Experimental source: clone vfa-ypt1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NKXD motif
F;151-153/Region: GTP-binding SAK/L motif
F;21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             43.3%;  Score 497.5;  DB 2;  Length 202;
  Best Local Similarity   45.8%;  Pred. No. 6.6e-36;
  Matches   93;  Conservative   46;  Mismatches   57;  Indels    7;  Gaps    3;

Qy          23 SPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQ 82
               :| ||   |::|:||:||||:| |::| | :::  ::|:|:|:||: : |  ||  :|||
Db           2 NPEYDYLFKLLLIGDSGVGKSCLLLRFADDSYID-SYISTIGVDFKIRTVEQDGKTIKLQ 60

Qy          83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLG 142
               |||||||||||::| :||| |  ::::||:|:: ||:|:: ||:||  ||  :|  :|:|
Db          61 IWDTAGQERFRTITSSYYRGAHGIIIVYDVTDEESFNNVKQWLSEIDRYASDNVNKLLVG 120

Qy         143 NKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQ---- 198
               || |::  | :  |  :  | | |:||:|||||   ||| ||:|:|  :| |   |
Db         121 NKCDLTENRAVPYETAKAFADEIGIPFMETSAKDSTNVEQAFMAMASSIKERMASQPTNN 180

Qy         199 ADEPSFQIRDYVESQKKRSSCCS 221
               |  |: |||     ||  | |||
Db         181 ARPPTVQIRGQPVGQK--SGCCS 201


Search completed: June 18, 2003, 15:49:08
Job time : 41 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 20:56:44 ; Search time 6785 Seconds
(without alignments)
11469.551 Million cell updates/sec

Title: US-09-817-199B-1
Perfect score: 2674
Sequence: 1 ttcgcctgcgggccggcact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2161.8	80.8	2195	9	AK057069	AK057069 Homo sapi
2	1962	73.4	1977	9	AK098068	AK098068 Homo sapi
3	1354.8	50.7	3105	9	AK054846	AK054846 Homo sapi
4	1086	40.6	1116	6	AX236082	AX236082 Sequence
5	938.4	35.1	1700	9	BC016615	BC016615 Homo sapi
6	576	21.5	576	6	AX236084	AX236084 Sequence
7	563.6	21.1	690	10	AF233582	AF233582 Mus muscu
8	315.6	11.8	1513	9	BC007681	BC007681 Homo sapi
9	312.4	11.7	1098	10	RNU18771	U18771 Rattus norv
10	304.6	11.4	1320	9	AB027137	AB027137 Homo sapi
11	303.2	11.3	573	9	AF498952	AF498952 Homo sapi
12	292.6	10.9	1340	6	AR062279	AR062279 Sequence
13	234.2	8.8	2536	3	AY061826	AY061826 Drosophil
c 14	214.8	8.0	127587	9	HS941F9	Z95331 Human DNA s
c 15	214.2	8.0	102332	9	AC010352	AC010352 Homo sapi
16	214.2	8.0	107465	9	AC010434	AC010434 Homo sapi
c 17	212.8	8.0	146740	9	CNS01DTX	AL132989 Human chr
c 18	212.6	8.0	174309	9	AC016656	AC016656 Homo sapi
19	212.6	8.0	177640	9	AC016652	AC016652 Homo sapi
c 20	211.6	7.9	148841	9	AC011462	AC011462 Homo sapi
c 21	210.4	7.9	199601	9	AL389915	AL389915 Human DNA
22	210.2	7.9	90000	9	AC117467	AC117467 Homo sapi
23	209.6	7.8	141771	9	AC091821	AC091821 Homo sapi
24	209	7.8	110000	2	AC013391_0	AC013391 Homo sapi
c 25	209	7.8	125980	9	AC093223	AC093223 Homo sapi
26	209	7.8	157141	9	AC011470	AC011470 Homo sapi
c 27	209	7.8	176304	9	AC013787	AC013787 Homo sapi
28	208.4	7.8	167509	2	AC012107	AC012107 Homo sapi
29	208	7.8	43790	9	AC113388	AC113388 Homo sapi
30	208	7.8	86685	9	AC118457	AC118457 Homo sapi
c 31	208	7.8	107995	2	HSJ356J11	AL121971 Homo sapi
c 32	208	7.8	172048	9	HS179N16	Z95152 Homo sapien
c 33	207.6	7.8	137271	9	AC004552	AC004552 Homo sapi
34	207.6	7.8	207366	9	AC090001	AC090001 Homo sapi
35	207.4	7.8	136222	9	AC004099	AC004099 Homo sapi
c 36	207.2	7.7	63905	9	AC022515	AC022515 Homo sapi
c 37	207.2	7.7	186330	9	AC007376	AC007376 Homo sapi
38	207.2	7.7	271814	2	AC027406	AC027406 Homo sapi
c 39	207	7.7	154289	2	AC027714	AC027714 Homo sapi
c 40	207	7.7	161757	2	AC044901	AC044901 Homo sapi
c 41	207	7.7	163528	2	AC130416	AC130416 Homo sapi
c 42	207	7.7	187399	9	AC090970	AC090970 Homo sapi
c 43	207	7.7	209797	2	AC005748	AC005748 Homo sapi
44	207	7.7	346940	2	AC107420	AC107420 Homo sapi
c 45	206.8	7.7	157113	9	AC083797	AC083797 Homo sapi

ALIGNMENTS

RESULT 1
AK057069
LOCUS AK057069 2195 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ32507 fis, clone SMINT1000048, moderately similar to Mus musculus GTPase Rab37 mRNA.
ACCESSION AK057069
VERSION AK057069.1 GI:16552648
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens small intestine cDNA to mRNA, clone_lib:SMINT1 clone:SMINT1000048.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,

Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2195)

AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.

TITLE Direct Submission

JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES Location/Qualifiers

source 1. .2195
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="SMINT1000048"
/tissue_type="small intestine"
/clone_lib="SMINT1"
/note="cloning vector: pME18SFL3"

BASE COUNT 511 a 596 c 582 g 506 t

ORIGIN

Query Match 80.8%; Score 2161.8; DB 9; Length 2195;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        472 AGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGG 90

Qy        532 AGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCT 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 AGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCT 150

Qy        592 TTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCT 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 TTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCT 210

Qy        652 TCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGT 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 TCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGT 270

Qy        712 GAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCC 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 GAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCC 330

Qy        772 AGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCAC 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 AGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCAC 390

Qy        832 AGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCG 450

Qy        892 CATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAA 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 CATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAA 510

Qy        952 AAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTA 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTA 570
```

```
Qy       1012 GGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAG 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 GGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAG 630

Qy       1072 CCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTC 690

Qy       1132 AGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAG 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 AGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAG 750

Qy       1192 CTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCC 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 CTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCC 810

Qy       1252 TCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTC 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTC 870

Qy       1312 TAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCA 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 TAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCA 930

Qy       1372 CACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAG 1431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 CACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAG 990

Qy       1432 ATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCA 1491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 ATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCA 1050

Qy       1492 CCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAG 1551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 CCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAG 1110

Qy       1552 TCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCA 1611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 TCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCA 1170

Qy       1612 CTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTG 1671
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1171 CTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCCCCGAAGTAGCTGGGACTATAGGTG 1230

Qy       1672 TGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTG 1731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 TGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTG 1290

Qy       1732 CCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTA 1791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 CCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTA 1350

Qy       1792 CTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTG 1851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 CTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTG 1410

Qy       1852 GCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGG 1911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 GCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGG 1470

Qy       1912 TTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAG 1971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 TTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAG 1530

Qy       1972 AGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATG 2031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 AGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATG 1590

Qy       2032 TATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTT 2091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 TATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTT 1650

Qy       2092 CTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGT 2151
```

```
                                                                                
Db       1651 CTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGT 1710

Qy       2152 GGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGT 2211
                                                                                
Db       1711 GGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGT 1770

Qy       2212 GGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATC 2271
                                                                                
Db       1771 GGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATC 1830

Qy       2272 TCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCA 2331
                                                                                
Db       1831 TCCCAAAGCTCCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCA 1890

Qy       2332 GGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGC 2391
                                                                                
Db       1891 GGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGC 1950

Qy       2392 TGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 2451
                                                                                
Db       1951 TGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 2010

Qy       2452 GCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGG 2511
                                                                                
Db       2011 GCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGG 2070

Qy       2512 AGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCC 2571
                                                                                
Db       2071 AGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCC 2130

Qy       2572 CATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGG 2631
                                                                                
Db       2131 CATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGG 2190

Qy       2632 AAAGC 2636
                 |||||
Db       2191 AAAGC 2195
```

```
RESULT 2
AK098068
LOCUS       AK098068                1977 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
            to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION   AK098068
VERSION     AK098068.1  GI:21758001
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens trachea cDNA to mRNA, clone_lib:TRACH2
            clone:TRACH2000540.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
            Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
            Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1977)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .1977
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TRACH2000540"
                     /tissue_type="trachea"
                     /clone_lib="TRACH2"
                     /note="cloning vector: pME18SFL3"
     CDS             57. .728
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC05227.1"
                     /db_xref="GI:21758002"
                     /translation="MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTC
                     FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRD
                     AQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGET
                     LAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSS
                     CCSFM"
BASE COUNT      442 a    560 c    539 g    436 t
ORIGIN
```

```
  Query Match             73.4%;  Score 1962;  DB 9;  Length 1977;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1962;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 TTCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCG 75

Qy         61 CCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 CCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACG 135

Qy        121 ACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 ACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGA 195

Qy        181 TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACT 255

Qy        241 TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 315

Qy        301 CTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 CTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGC 375

Qy        361 TTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTG 435

Qy        421 AGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATA 495

Qy        481 TGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTG 555

Qy        541 TTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA 615

Qy        601 TCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        616 TCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCC 675

Qy        661 GAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAG 735

Qy        721 GGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGC 795

Qy        781 TTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 TTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 855

Qy        841 CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCCGCATCCACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCCGCATCCACAG 915

Qy        901 GGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGC 975

Qy        961 CTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 CTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGG 1035

Qy       1021 GGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 GGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGG 1095

Qy       1081 ACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 ACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCT 1155

Qy       1141 GGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 GGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTC 1215

Qy       1201 GCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1216 GCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAG 1275

Qy       1261 ATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 ATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCT 1335

Qy       1321 ATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 ATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCAC 1395

Qy       1381 ATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 ATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGG 1455

Qy       1441 GAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 GAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAAC 1515

Qy       1501 CTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 CTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCT 1575

Qy       1561 GTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 GTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGG 1635

Qy       1621 GCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 GCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATC 1695

Qy       1681 ACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 ACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTG 1755

Qy       1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 GTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTA 1815

Qy       1801 CACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 CACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATG 1875

Qy       1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCC 1935

Qy       1921 AGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTG 1962
              |||||||||||||||||||||||||||||||||||||||||||
Db       1936 AGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTG 1977
```

RESULT 3
AK054846

```
LOCUS       AK054846               3105 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately
            similar to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION   AK054846
VERSION     AK054846.1  GI:16549463
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
            clone:BRACE2002812.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3105)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .3105
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BRACE2002812"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE2"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      680 a    905 c    898 g    622 t
ORIGIN
```

```
  Query Match             50.7%;  Score 1354.8;  DB 9;  Length 3105;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1362;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        589 CCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCA 648
              || || |  |     | | ||||||||||||||||||||||||||||||||||||||||||
Db       1732 CCATTGTCTCTTCTTCAAGGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCA 1791
```

```
Qy       649 GCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCA 708
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1792 GCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCA 1851

Qy       709 TGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCT 768
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1852 TGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCT 1911

Qy       769 CCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTG 828
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1912 CCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTG 1971

Qy       829 CACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCC 888
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1972 CACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCC 2031

Qy       889 CCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATTAGTACATAATTTAATAC 948
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2032 CCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATTAGTACATAATTTAATAC 2091

Qy       949 CAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 1008
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2092 CAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 2151

Qy      1009 CTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCT 1068
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2152 CTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCT 2211

Qy      1069 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACC 1128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2212 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACC 2271

Qy      1129 CTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAG 1188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2272 CTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAG 2331

Qy      1189 GAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAA 1248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2332 GAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAA 2391

Qy      1249 GCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGG 1308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2392 GCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGG 2451

Qy      1309 GTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCA 1368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2452 GTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCA 2511

Qy      1369 GCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 1428
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2512 GCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 2571

Qy      1429 CAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 1488
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2572 CAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 2631

Qy      1489 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATG 1548
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2632 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATG 2691

Qy      1549 GAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCT 1608
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2692 GAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCT 2751

Qy      1609 CCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAG 1668
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2752 CCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCCCCGAAGTAGCTGGGACTATAG 2811

Qy      1669 GTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATG 1728
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2812 GTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATG 2871
```

```
Qy      1729 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAA 1788
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2872 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAA 2931

Qy      1789 GTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCT 1848
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2932 GTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCT 2991

Qy      1849 TTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 1908
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2992 TTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 3051

Qy      1909 GGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTG 1962
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3052 GGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTG 3105
```

RESULT 4
AX236082
LOCUS AX236082 1116 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 13 from Patent WO0164887.
ACCESSION AX236082
VERSION AX236082.1 GI:15795889
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1116)
AUTHORS Meyers,R.A.
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins
JOURNAL Patent: WO 0164887-A 13 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. .1116
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 124. .699
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC88255.1"
/db_xref="GI:15795890"
/translation="MLLGDTGVGKTCFLIQFKDGAFLSGTFIATVGIDFRNKVVTVDG
VRVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQR
DVVIMLLGNKADMSSERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFLAIAKELK
YRAGHQADEPSFQIRDYVESQKKRSSCCSFM"
BASE COUNT 260 a 313 c 322 g 221 t
ORIGIN

Query Match 40.6%; Score 1086; DB 6; Length 1116;
Best Local Similarity 99.5%; Pred. No. 5.1e-280;
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         37 GGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 96
              | | |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 GCGTCCGGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 82

Qy         97 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 142

Qy        157 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 202

Qy        217 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 262

Qy        277 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 322
```

```
Qy        337 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 382

Qy        397 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 442

Qy        457 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 502

Qy        517 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 562

Qy        577 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 622

Qy        637 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 682

Qy        697 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 742

Qy        757 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 802

Qy        817 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 862

Qy        877 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 922

Qy        937 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 982

Qy        997 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1042

Qy       1057 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGA 1116
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1043 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAAGTAATGA 1102

Qy       1117 CTCCATCGCACCCT 1130
              ||||||||||||||
Db       1103 CTCCATCGCACCCT 1116
```

```
RESULT 5
BC016615
LOCUS       BC016615               1700 bp    mRNA    linear   PRI 05-NOV-2001
DEFINITION  Homo sapiens, Similar to RAB37, member of RAS oncogene family,
            clone MGC:21391 IMAGE:4520191, mRNA, complete cds.
ACCESSION   BC016615
VERSION     BC016615.1  GI:16741620
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1700)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 28 Row: j Column: 15
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES              Location/Qualifiers
     source           1. .1700
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="MGC:21391 IMAGE:4520191"
                      /tissue_type="Prostate, adenocarcinoma."
                      /clone_lib="NIH_MGC_91"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              3. .674
                      /codon_start=1
                      /product="Similar to RAB37, member of RAS oncogene family"
                      /protein_id="AAH16615.1"
                      /db_xref="GI:16741621"
                      /translation="MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTC
                      FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFRSVTHAYYRD
                      AQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSSERVIRSEDGET
                      LAREYGVPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVESQKKRSS
                      CCSFM"
BASE COUNT      437 a    444 c    460 g    359 t
ORIGIN
```

```
  Query Match           35.1%;  Score 938.4;  DB 9;  Length 1700;
  Best Local Similarity  99.9%;  Pred. No. 2.2e-240;
  Matches  939;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1735 AGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTG 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 AGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTG 811

Qy       1795 GGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCA 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 GGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCA 871

Qy       1855 GGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTA 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        872 GGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTA 931

Qy       1915 AGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAGA 1974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        932 AGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAGA 991

Qy       1975 AATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTAT 2034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        992 AATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTAT 1051

Qy       2035 CCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTC 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 CCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTC 1111
```

```
Qy       2095 AAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGT 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1112 AAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGT 1171

Qy       2155 GGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGG 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1172 GGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGG 1231

Qy       2215 TATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCC 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 TATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCC 1291

Qy       2275 CAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGC 2334
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1292 CAAAGCTCCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGC 1351

Qy       2335 ATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGC 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1352 ATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGC 1411

Qy       2395 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1412 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCG 1471

Qy       2455 GCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1472 GCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGG 1531

Qy       2515 AGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCAT 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1532 AGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCAT 1591

Qy       2575 AGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGGAAA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1592 AGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGGAAA 1651

Qy       2635 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
             ||||||||||||||||||||||||||||||||||||||||
Db       1652 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1691
```

```
RESULT 6
AX236084
LOCUS       AX236084                576 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 15 from Patent WO0164887.
ACCESSION   AX236084
VERSION     AX236084.1  GI:15795891
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Meyers,R.A.
  TITLE     32705, 23224, 27423, 32700, 32712, novel human g-proteins
  JOURNAL   Patent: WO 0164887-A 15 07-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      141 a    148 c    167 g    120 t
ORIGIN

  Query Match             21.5%;  Score 576;  DB 6;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-143;
  Matches  576;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60
```

```
Qy        198 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy        258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy        318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        378 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        438 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        498 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        558 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        618 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        678 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 713
              ||||||||||||||||||||||||||||||||||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 576
```

```
RESULT 7
AF233582
LOCUS       AF233582                690 bp    mRNA    linear   ROD 02-MAY-2000
DEFINITION  Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
ACCESSION   AF233582
VERSION     AF233582.1  GI:7677421
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   Masuda,E.S., Luo,Y., Young,C., Shen,M., Rossi,A.B., Huang,B.C.,
            Yu,S., Bennett,M.K., Payan,D.G. and Scheller,R.H.
  TITLE     Rab37 is a novel mast cell specific GTPase localized to secretory
            granules
  JOURNAL   FEBS Lett. 470 (1), 61-64 (2000)
  MEDLINE   20189834
   PUBMED   10722846
REFERENCE   2  (bases 1 to 690)
  AUTHORS   Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
            Avenue, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .690
                     /gene="Rab37"
     CDS             1. .672
                     /gene="Rab37"
                     /codon_start=1
                     /product="GTPase Rab37"
                     /protein_id="AAF67162.1"
                     /db_xref="GI:7677422"
```

```
                 /translation="MTGTPGAATAGDGEAPERSPPFSPNYDLTGKVMLLGDSGVGKTC
                 FLIQFKDGAFLSGTFIATVGIDFRNKVVTVDGARVKLQIWDTAGQERFRSVTHAYYRD
                 AQALLLLYDITNQSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADVSSERVIRSEDGET
                 LAREYGVPFMETSAKTGMNVELAFLAIAKELKYRAGRQPDEPSFQIRDYVESQKKRSS
                 CCSFV"
BASE COUNT      160 a    181 c    201 g    148 t
ORIGIN

  Query Match           21.1%;  Score 563.6;  DB 10;  Length 690;
  Best Local Similarity  88.6%;  Pred. No. 5.9e-140;
  Matches  611;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy         42 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 101
              ||||| ||||| ||||| || | | || |  ||||||||||||||| ||||||||||||
Db          1 ATGACTGGCACACCAGGAGCTGCTACCGCTGGGGATGGCGAGGCCCCTGAGCGCTCCCCG 60

Qy        102 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 161
              |||| ||| |||| |||||| ||||| |||||||||||||||||| |||||| | ||||||
Db         61 CCCTTCAGCCCGAACTACGATCTCACCGGCAAGGTGATGCTTCTTGGAGACTCGGGCGTC 120

Qy        162 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 221
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 180

Qy        222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 281
              |||||||||||||||||||||||| || |||||||| |||||||| |  || ||||||
Db        181 GCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTGCCAGGGTGAAG 240

Qy        282 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 341
              || |||||||||||| || || ||||| || ||||| || || ||||||||||||||||
Db        241 CTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCCGCAGTGTGACCCATGCTTATTAC 300

Qy        342 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 401
              |||||||||||| ||||| ||| |||||||||||||||| | || ||||| ||||||
Db        301 CGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATCACCAACCAGTCCTCTTTTGACAAC 360

Qy        402 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 461
              |||||||||||||||| ||||||||||||||||||||||||||||||||||| |||||
Db        361 ATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATTATGCTT 420

Qy        462 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 521
              ||||||||||||| ||| | |||||||||| |||||||||| ||||| ||||||||
Db        421 CTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACA 480

Qy        522 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 581
              |||||||||| || |||||||| ||| |||||||||| |||||||||||||||||| |||
Db        481 CTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTG 540

Qy        582 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 641
              ||||| |||||||||| || |||||||||||||||||||| || |||   ||| | |||
Db        541 GAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTGAT 600

Qy        642 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        601 GAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 660

Qy        702 TCCTTCATGTGAATCCCAGGGGGCAGAGAG 731
              ||||| ||||| ||| ||||| ||||
Db        661 TCCTTTGTGTGACCCCCTAGGGGCTAAGAG 690


RESULT 8
BC007681
LOCUS       BC007681                1513 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, RAB26, member RAS oncogene family, clone MGC:3503
            IMAGE:3627067, mRNA, complete cds.
ACCESSION   BC007681
VERSION     BC007681.1  GI:14043378
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1  (bases 1 to 1513)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 12 Row: m Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 5931611.
FEATURES             Location/Qualifiers
     source          1. .1513
                     /organism="Homo sapiens"
                     /db_xref="LocusID:25837"
                     /db_xref="taxon:9606"
                     /clone="MGC:3503 IMAGE:3627067"
                     /tissue_type="Uterus, endometrium adenocarcinoma"
                     /clone_lib="NIH_MGC_44"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             187. .759
                     /codon_start=1
                     /product="RAB26, member RAS oncogene family"
                     /protein_id="AAH07681.1"
                     /db_xref="GI:14043379"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      326 a    452 c    449 g    286 t
ORIGIN

  Query Match           11.8%;  Score 315.6;  DB 9;  Length 1513;
  Best Local Similarity  71.2%;  Pred. No. 2.4e-73;
  Matches  417;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              |||||||  || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db        165 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 224

Qy        176 CCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCAT 235
              ||| | | |||||| || || || |||||| | || ||||||||  ||||||| |||||
Db        225 GCTGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCAT 284

Qy        236 AGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGA 295
              |||||| |||||||  ||   ||  ||||||||  ||||  ||||||||||||||  |||||
Db        285 TGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGA 344

Qy        296 CACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGC 355
              ||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| ||
Db        345 CACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGC 404
```

```
Qy        356 CTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCT 415
              |||| ||||| || ||  |||||||||  | || || |||||||||  |||||||||
Db        405 TCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCT 464

Qy        416 CACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGC 475
              || ||||| || ||||| ||||||   ||||||| | ||||||||||| || |||||||
Db        465 GACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGT 524

Qy        476 GGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTA 535
              |||       |   ||  | ||| |        || ||||| |||   |||||| ||||||
Db        525 GGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTA 584

Qy        536 CGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCT 595
              ||  | |||||| ||||||||||||||||||||| ||| | || ||||| || |||||
Db        585 TGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCAC 644

Qy        596 GGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCA 655
              ||||| || |||||  ||||  | ||  ||  | |     |    |||||  ||||||
Db        645 AGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCG 704

Qy        656 GATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              | | |  || || ||  ||    ||       | ||  |||||||
Db        705 GCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 750
```

RESULT 9
RNU18771
LOCUS RNU18771 1098 bp mRNA linear ROD 18-JUL-1995
DEFINITION Rattus norvegicus Rab26 mRNA, complete cds.
ACCESSION U18771
VERSION U18771.1 GI:619733
KEYWORDS .
SOURCE rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1098)
AUTHORS Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
TITLE Molecular cloning of a new member of the Rab protein family, Rab 26, from rat pancreas
JOURNAL Biochem. Biophys. Res. Commun. 207 (3), 950-956 (1995)
MEDLINE 95169156
PUBMED 7864900
REFERENCE 2 (bases 1 to 1098)
AUTHORS Williams,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA
FEATURES Location/Qualifiers
source 1. .1098
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="pancreas"
/dev_stage="adult"
5'UTR 1. .29
CDS 30. .602
/codon_start=1
/product="Rab26"
/protein_id="AAA69955.1"
/db_xref="GI:619734"
/translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
MKVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDITNKDSFDNIQAWLTEIQEYAQQ
DVVLMLLGNKVDSTQERVVKREDGEKLAKEYGLPFMETSAKSGLNVDLAFTAIAKELK
QRSTKAPSEPRFRLHDYVKREGRGVSCCRL"
3'UTR 603. .1098
polyA_signal 1063. .1068
BASE COUNT 267 a 277 c 306 g 248 t
ORIGIN

Query Match 11.7%; Score 312.4; DB 10; Length 1098;
Best Local Similarity 71.8%; Pred. No. 1.6e-72;
Matches 409; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

```
Qy        131 CAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAA 190
              |||||| |||||  |||| ||  | || || || || || ||  | ||  | |  |||||
Db         23 CAAGGTCATGCTGGTGGGGGATTCCGGTGTGGGGAAGACGTGCCTGCTCGTGCGCTTCAA 82

Qy        191 AGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAA 250
              || |||||||||||| | || ||||||||  |||| || |||||  |||||  |||| ||
Db         83 GGATGGGGCCTTCCTGGCTGGTACCTTCATCTCCACTGTGGGCATCGACTTCCGGAATAA 142

Qy        251 GGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGA 310
              ||  ||  |||||||||| |||  |||||||||||||||||||||||| ||||| ||||||
Db        143 AGTTCTGGATGTGGATGGCATGAAGGTGAAGCTGCAGATCTGGGACACGGCTGGTCAGGA 202

Qy        311 ACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTA 370
              |||||||||| |||||||||| || ||| | |||||||| ||  || | |||| ||
Db        203 GCGGTTCCGAAGTGTCACCCATGCCTACTACCGGGATGCTCACGCACTGTTGCTGCTCTA 262

Qy        371 TGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGA 430
              |||||||||||||||  ||| ||||||||||||  |||||| | || || || || || ||
Db        263 CGACATCACCAACAAAGATTCCTTCGACAACATCCAGGCCTGGTTGACAGAAATCCAGGA 322

Qy        431 GTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGA 490
              |||||||||  ||||||||||  |||||||| | || |||||||  ||    |    ||
Db        323 ATATGCCCAGCAGGACGTGGTGCTCATGCTGTTGGGGAACAAGGTTGACTCTACTCAAGA 382

Qy        491 AAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCT 550
              | |||| |        |||| || ||||  || |||| ||||||  ||  | || ||| |
Db        383 ACGAGTGGTAAAGAGGGAAGATGGGGAGAAATTAGCCAAGGAGTATGGGCTGCCATTCAT 442

Qy        551 GGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGA 610
              ||||||||||||||||  ||| | |||||||| || || |||   |||| || ||||
Db        443 GGAGACCAGCGCCAAGAGCGGCCTCAATGTGGACTTGGCTTTTACAGCCATAGCAAAGGA 502

Qy        611 ACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGT 670
              |||||| |  |  ||     |    |  ||||||  ||||  | | |  |||||||
Db        503 GCTGAAACAAAGATCCACCAAGGCTCCCAGTGAGCCGCGCTTCAGGCTGCATGACTATGT 562

Qy        671 AGAGTCCCAGAAGAAGCGCTCCAGCTGCTG 700
                ||    ||      |   |  ||||||
Db        563 GAAGAGGGAGGGCCGAGGGGTCTCCTGCTG 592
```

RESULT 10
AB027137
LOCUS AB027137 1320 bp mRNA linear PRI 29-SEP-1999
DEFINITION Homo sapiens v46133 mRNA for RAB-26, complete cds.
ACCESSION AB027137
VERSION AB027137.1 GI:5931611
KEYWORDS RAB-26; v46133.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S., Muramatsu,M. and Saito,T.
TITLE Human RAS-related protein RAB-26
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S., Muramatsu,M. and Saito,T.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1999) Toshiyuki Saito, National Institute of Radiological Sciences, Genome Research Group; Inage-ku Anagawa 4-9-1, Chiba, Chiba 263-8555, Japan (E-mail:t_saito@nirs.go.jp, Tel:81-43-201-3135, Fax:81-43-251-9818)
FEATURES Location/Qualifiers
source 1. .1320
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .1320

```
                     /gene="v46133"
     CDS             4. .576
                     /gene="v46133"
                     /note="RAS-related protein"
                     /codon_start=1
                     /product="RAB-26"
                     /protein_id="BAA84707.1"
                     /db_xref="GI:5931612"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      273 a    379 c    395 g    273 t
ORIGIN

  Query Match           11.4%;  Score 304.6;  DB 9;  Length 1320;
  Best Local Similarity   71.1%;  Pred. No. 2e-70;
  Matches  403;  Conservative    0;  Mismatches  164;  Indels    0;  Gaps    0;

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194
              || |||||  |||| ||| | || || || || || ||| | ||| | | |||||| ||
Db          1 GTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTGCTGGTGCGATTCAAGGAT 60

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
              || || |||||| | || |||||||  ||||||| |||| |||||| ||||||| ||
Db         61 GGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAAGTT 120

Qy        255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
                ||   |||||||| |||| |||||||||||||| |||||||| ||||| ||||| |||
Db        121 CTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGGAGCGG 180

Qy        315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
              ||||| || || |||||||| || ||| | |||||||| ||  |||| ||||| || ||
Db        181 TTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCTCTGCTGCTGCTCTACGAT 240

Qy        375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
              ||||||||||  | || || |||||||||  |||||||| || ||||| || |||||
Db        241 GTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTGACCGAGATCCACGAGTAC 300

Qy        435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 494
              ||||||   ||||||| | |||||||||| || ||||||| |||       |   || |
Db        301 GCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGT 360

Qy        495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 554
              ||| |       || ||||| ||||   |||||| |||||| ||  | ||||||  |||||
Db        361 GTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTATGGACTGCCCTTCATGGAG 420

Qy        555 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 614
              |||||||||||||| ||| | || ||||| || |||||    ||||| || |||||  ||
Db        421 ACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCACAGCCATAGCAAAGGAGTTG 480

Qy        615 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 674
              ||  | ||  ||  | |      |    |||||  |||||| | | |  || || ||  ||
Db        481 AAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCGGCTGCATGATTACGTTAAG 540

Qy        675 TCCCAGAAGAAGCGCTCCAGCTGCTGC 701
                ||       |  ||  |||||||
Db        541 AGGGAGGGTCGAGGGGCCTCCTGCTGC 567


RESULT 11
AF498952
LOCUS       AF498952                573 bp    mRNA    linear   PRI 01-MAY-2002
DEFINITION  Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete
            cds.
ACCESSION   AF498952
VERSION     AF498952.1  GI:20379079
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
  TITLE     Homo sapiens RAB family small GTP binding protein RAB26
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 573)
  AUTHORS   Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
            Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="brain"
     gene            1. .573
                     /gene="RAB26"
     CDS             1. .573
                     /gene="RAB26"
                     /note="RAB family member"
                     /codon_start=1
                     /product="small GTP binding protein RAB26"
                     /protein_id="AAM21100.1"
                     /db_xref="GI:20379080"
                     /translation="MLVGDSGVGKTCLLVRFKDGAFLAGTFISTVGIDFRNKVLDVDG
                     VKVKLQMWDTAGQERFRSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLTEIHEYAQH
                     DVALMLLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFTAIAKELK
                     QRSMKAPSEPRFRLHDYVKREGRGASCCRP"
BASE COUNT      117 a    151 c    191 g    114 t
ORIGIN

  Query Match           11.3%;  Score 303.2;  DB 9;  Length 573;
  Best Local Similarity   71.1%;  Pred. No. 4e-70;
  Matches  401;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy        138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 197
              |||||  |||| ||| | || || || || || ||| | ||| | | |||||| || ||
Db          1 ATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCTGCTGGTGCGATTCAAGGATGGT 60

Qy        198 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 257
              || |||||| | || |||||||  ||||||| |||| |||||| ||||||| ||  ||
Db         61 GCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAAGTTCTG 120

Qy        258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 317
                |||||||| ||||  |||||||||||||| |||||||| ||||| |||||| ||||||
Db        121 GACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGGAGCGGTTC 180

Qy        318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 377
              || || || |||||||| || ||| | ||||||| ||  |||| ||||| || ||  ||
Db        181 CGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGCTCTGCTGCTGCTCTACGATGTC 240

Qy        378 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 437
              ||||||||  | || || |||||||||  |||||||| || ||||| || ||||| |||
Db        241 ACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCTGACCGAGATCCACGAGTACGCC 300

Qy        438 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 497
              |||   ||||||| | |||||||||| || |||||||  |||       |   ||  | |||
Db        301 CAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGTGTG 360

Qy        498 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 557
              |       || ||||| ||||   |||||| |||||| ||  | ||||||  ||||||||
Db        361 GTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGTATGGACTGCCCTTCATGGAGACC 420

Qy        558 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 617
              |||||||||||  ||| | || ||||| || |||||    ||||| || |||||  ||||
Db        421 AGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCACAGCCATAGCAAAGGAGTTGAAG 480

Qy        618 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
              | ||  ||  | |      |    |||||  |||||| | | |  || || ||  ||
Db        481 CAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCCGGCTGCATGATTACGTTAAGAGG 540

Qy        678 CAGAAGAAGCGCTCCAGCTGCTGC 701
                ||       |  ||  |||||||
Db        541 GAGGGTCGAGGGGCCTCCTGCTGC 564
```

```
RESULT 12
AR062279
LOCUS       AR062279               1340 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5843717.
ACCESSION   AR062279
VERSION     AR062279.1  GI:5989970
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1340)
  AUTHORS   Hillman,J.L. and Guegler,K.J.
  TITLE     Rab protein
  JOURNAL   Patent: US 5843717-A 2 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .1340
                     /organism="unknown"
BASE COUNT      276 a    387 c    403 g    274 t
ORIGIN

  Query Match           10.9%;  Score 292.6;  DB 6;  Length 1340;
  Best Local Similarity   70.9%;  Pred. No. 3.4e-67;
  Matches  416;  Conservative    0;  Mismatches  169;  Indels    2;  Gaps    2;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              ||||||| || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db         21 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 80

Qy        176 CCTGA-TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCA 234
               |||  | | |||||| || || || |||||| | || |||||||| |||||||  |||
Db         81 GCTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139

Qy        235 TAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGG 294
              | |||||| ||||||| ||  ||   |||||||| |||| |||||||||||||| ||||
Db        140 TTGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGG 199

Qy        295 ACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGG 354
              |||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| |
Db        200 ACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATG 259

Qy        355 CCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGC 414
              |  |||| ||||| || ||  |||||||||| | || || |||||||||  ||||||||
Db        260 CTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGC 319

Qy        415 TCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGG 474
              | || ||||| || ||||| ||||||   ||||||| | |||||||||| || |||||||
Db        320 TGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGG 379

Qy        475 CGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGT 534
               |||      |   ||  | ||| |      || ||||| ||||   |||||| |||||
Db        380 TGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGT 439

Qy        535 ACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTC 594
              |  ||  | |||||| |||||||||||||||||||| ||| | || ||||| || |||||
Db        440 ATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCA 499

Qy        595 TGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCC 654
               ||||| || |||||  ||||  | ||  ||  | |      |    |||||  ||||||
Db        500 CAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCC 559

Qy        655 AGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
               | | |  || || ||  ||    ||       |  ||  |||||||
Db        560 GGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606


RESULT 13
AY061826
LOCUS       AY061826               2536 bp    mRNA    linear   INV 10-NOV-2001
DEFINITION  Drosophila melanogaster GH21984 full length cDNA.
ACCESSION   AY061826
VERSION     AY061826.1  GI:16902019
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 2536)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .2536
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="78C9-78D1"
     gene            1. .2536
                     /gene="CG7605"
                     /note="alignment with genomic scaffold AE003594"
                     /db_xref="FLYBASE:FBgn0037072"
     CDS             1172. .2338
                     /gene="CG7605"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="GH21984p"
                     /protein_id="AAL27637.1"
                     /db_xref="GI:16902020"
                     /db_xref="FLYBASE:FBgn0037072"
                     /translation="MASTAVGLGGGEGDPGAGGPPAGSAHPDDASSMSDDVFEDAETT
                     QARIEELRRRPFGDGSYNPPAAPASVSASITTTTTQQQQQHHNPSHHHQSSHHQPSHH
                     HHHHHHSQLSLTGSHHYHDDAIMAPVQRSATGYPGYRPSREAMQMYAYGTDDYDDDYN
                     DGWRSYRYDEVDMHPAPSNAHQQPFDDTVNHKTILLGDSGVGKTSFLVKYNTGEFRLG
                     SFSATVGIALTNKVVVVDGTRVKLQIWDTAGQERFRSVTHAYYRDAHALLLLYDVTNK
                     TTYDNIRAWLGEIREYAQEDVVIVLIGNKADCSGSERQVKREDGERLGREHNVPFMET
                     SAKTGLNVELSFTAVARQLKSRGYEHGDDGRFNVHDFVRDNTKARSVCAQCRNM"
BASE COUNT      831 a    621 c    624 g    460 t
ORIGIN

  Query Match           8.8%;  Score 234.2;  DB 3;  Length 2536;
  Best Local Similarity   65.7%;  Pred. No. 1.8e-51;
  Matches  357;  Conservative    0;  Mismatches  183;  Indels    3;  Gaps    1;

Qy        131 CAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAA 190
              ||||  ||| || ||||| ||  | ||||| || || || |  |||||| || ||| |||
Db       1744 CAAGACGATCCTGCTGGGGGATTCGGGCGTGGGAAAGACCTCCTTCCTGGTCAAATACAA 1803

Qy        191 AGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAA 250
                 || |  ||||     ||  ||||    |||||·|| ||||| |   | | ||||||
Db       1804 TACGGGCGAGTTCCGACTGGGTTCCTTTTCCGCCACAGTGGGCATTGCGCTTACGAACAA 1863
```

```
Qy        251 GGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGA 310
              ||||||   || |||||   | | || |||||||| |||||||||| ||||| |||||
Db       1864 AGTGGTGGTCGTCGATGGAACGCGCGTCAAGCTGCAAATCTGGGACACAGCTGGTCAGGA 1923

Qy        311 ACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTA 370
              || ||||| ||||| ||||| || |||||  | || || || ||  | || ||||||||
Db       1924 GCGATTCCGGAGCGTTACCCACGCCTATTATCGCGACGCGCACGCTCTACTGCTGCTGTA 1983

Qy        371 TGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGA 430
              ||| | ||||||||  |  | |  ||||||||  | ||||||||    ||||| |  ||
Db       1984 CGACGTGACCAACAAGACCACCTATGACAACATTCGCGCCTGGCTGGGCGAGATCCGGGA 2043

Qy        431 GTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGA---GCAG 487
              ||| || |||  ||||||||| ||| |  |  |||||||||||| ||    |  ||||
Db       2044 GTACGCGCAGGAGGACGTGGTCATCGTTTTAATAGGCAACAAGGCCGACTGCAGCGGCAG 2103

Qy        488 CGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTT 547
              |||  |  | |       || || || |||   ||||   ||||| ||   || |||||
Db       2104 CGAGCGGCAGGTGAAGCGGGAGGATGGGGAGCGTTTGGGGCGGGAGCACAACGTGCCCTT 2163

Qy        548 CCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAA 607
              | ||||||||   |||||||| ||  | |||||||| |  ||||    ||  | ||||
Db       2164 CATGGAGACCTCGGCCAAGACGGGACTCAATGTGGAGCTGTCCTTCACAGCGGTGGCCAG 2223

Qy        608 GGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTA 667
              | |||| ||    || | |    | || |  |||||    || ||| |  | |  || |
Db       2224 GCAACTAAAGAGTCGCGGCTACGAGCACGGCGATGATGGAAGGTTCAATGTGCATGATTT 2283

Qy        668 TGT 670
              |||
Db       2284 TGT 2286
```

RESULT 14
HS941F9/c
LOCUS HS941F9 127587 bp DNA linear PRI 12-DEC-1999
DEFINITION Human DNA sequence from clone CTA-941F9 on chromosome 22q13 Contains the 3' end of the FBLN1 gene for Fibulin 1 isoforms B, C and D, the first exon of the gene for a novel protein (the ortholog of mouse brain protein E46), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.
ACCESSION Z95331
VERSION Z95331.2 GI:6572282
KEYWORDS HTG; E46; FBLN1; Fibulin 1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 127587)
AUTHORS Connor,R.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Dec 13, 1999 this sequence version replaced gi:2281936.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr22
CTA-941F9 is from the human BAC library described in U-J. Kim et al. (1996) Genomics 34, 213-218.
VECTOR: pBeloBAC11
This sequence is the entire insert of clone CTA-941F9 The true left end of clone RP3-398C22 is at 87881 in this sequence. The true left end of clone LL22NC03-38E2 is at 9039 in this sequence. The true right end of clone LL22NC03-38E2 is at 54042 in this sequence.

FEATURES Location/Qualifiers

source	1. .127587
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="22"
	/map="q13"
	/clone="CTA-941F9"
	/clone_lib="CIT978SK-A2"
misc_feature	18. .560
	/note="match: GSS: Em:B14132"
misc_feature	25. .461
	/note="match: GSS: Em:AQ777382"
misc_feature	complement(1107. .1512)
	/note="match: GSS: Em:AQ060472"
repeat_region	1436. .1467
	/note="16 copies 2 mer gt 84 conserved"
repeat_region	2424. .2647
	/note="7 copies 32 mer 94 conserved"
repeat_region	2664. .2735
	/note="MLT2B repeat: matches 381. .448 of consensus"
repeat_region	2764. .3168
	/note="MLT2B repeat: matches 1. .409 of consensus"
repeat_region	3169. .3439
	/note="HERVL repeat: matches 5384. .5654 of consensus"
repeat_region	3440. .3752
	/note="AluSx repeat: matches 1. .311 of consensus"
repeat_region	3753. .9139
	/note="HERVL repeat: matches 1. .5384 of consensus"
repeat_region	9140. .9577
	/note="MLT2B repeat: matches 1. .444 of consensus"
repeat_region	9778. .9851
	/note="MIR repeat: matches 73. .141 of consensus"
gene	10369. .48592
	/gene="FBLN1"
mRNA	<10369. .10819
	/gene="FBLN1"
	/product="bK941F9.2 (Fibulin 1 (isoform C))"
	/note="match: cDNAs: Em:AF013752 Em:AF051400 Em:X53743 Em:X70853 Em:AL050095 Em:X75285 Em:X82494; match: ESTs: Em:R25317 Em:AA226244 Em:AV064777 Em:R62365 Em:AA508812 Em:AA041958 Em:AA626088 Em:AI092723 Em:AI130996 Em:AI188632 Em:AI273494 Em:H02670 Em:AI893154 Em:AI167845 Em:AI690847 Em:AA368999 Em:R64034 Em:F16897 Em:AA142940 Em:AV064836 Em:R23566 Em:AA010959 Em:R76273 Em:R68191 Em:R68138 Em:C02271 Em:AA555258 Em:AI066730 Em:AI087421 Em:AI493361 Em:AI803546 Em:AI828352 Em:C18390 Em:AA781778 Em:R12400 Em:T49888 Em:AI084677 Em:AI553895 Em:AI858475"
	/evidence=not_experimental
CDS	<10369. .10723
	/gene="FBLN1"
	/note="match: proteins: Sw:P23144 Sw:Q08878"
	/codon_start=2
	/evidence=not_experimental
	/product="bK941F9.2 (Fibulin 1 (isoform C))"
	/protein_id="CAB62960.1"
	/db_xref="GI:6572283"
	/translation="RCERLPCHENRECSKLPLRITYYHLSFPTNIQAPAVVFRMGPSS AVPGDSMQLAITGGNEEGFFTTRKVSPHSGVVALTKPVPEPRDLLLTVKMDLSRHGTV SSFVAKLFIFVSAEL"
misc_feature	10374. .10803
	/gene="FBLN1"
	/note="match: STS: Em:R76273"

```
polyA_signal    complement(10783. .10788)
polyA_site      10813
                /gene="FBLN1"
polyA_site      10819
                /gene="FBLN1"
repeat_region   11114. .11149
                /note="MIR repeat: matches 109. .144 of consensus"
repeat_region   11527. .11572
                /note="MIR repeat: matches 75. .120 of consensus"
repeat_region   11651. .11934
                /note="L2 repeat: matches 1992. .2328 of consensus"
repeat_region   12001. .12270
                /note="L2 repeat: matches 2454. .2750 of consensus"
misc_feature    12337. .13158
                /gene="FBLN1"
                /note="match: STS: Em:X53742"
mRNA            <12345. .13158
                /gene="FBLN1"
                /product="bK941F9.2 (Fibulin 1 (isoform B))"
                /note="match: cDNAs: Em:AB007943 Em:AF078528 Em:Y14610
                Em:AF011357 Em:X53742 Em:AF019887; match: ESTs: Em:R78704
                Em:R33281 Em:R32733 Em:H01971 Em:H02060"
                /evidence=not_experimental
CDS             <12345. .12449
                /gene="FBLN1"
                /note="match: proteins: Sw:P23143"
                /codon_start=1
                /evidence=not_experimental
                /product="bK941F9.2 (Fibulin 1 (isoform B))"
                /protein_id="CAB62961.1"
                /db_xref="GI:6572284"
                /translation="KSKKGRQNTPAGSSKEDCRVLPWKQGLEDTHLDA"
misc_feature    12449. .12703
                /gene="FBLN1"
                /note="match: STS: Em:G59664"
misc_feature    12450. .12849
                /gene="FBLN1"
                /note="match: STS: Em:G11168"
misc_feature    complement(12733. .13139)
                /note="match: STS: Em:H01971"
repeat_region   12895. .13143
                /note="L2 repeat: matches 2419. .2705 of consensus"
misc_feature    complement(12974. .13280)
                /note="match: STS: Em:G08093; match: STS: Em:G08093"
polyA_signal    13138. .13143
                /gene="FBLN1"
polyA_site      13158
                /gene="FBLN1"
repeat_region   13237. .13536
                /note="AluSp repeat: matches 1. .300 of consensus"
repeat_region   13974. .14045
                /note="MIR repeat: matches 50. .122 of consensus"
repeat_region   14352. .14607
                /note="MIR repeat: matches 2. .262 of consensus"
repeat_region   14791. .15080
                /note="AluSx repeat: matches 1. .290 of consensus"
repeat_region   15714. .15856
                /note="MLT1C repeat: matches 300. .451 of consensus"
repeat_region   16332. .16625
                /note="MLT1C repeat: matches 1. .310 of consensus"
repeat_region   16637. .16815
                /note="MLT1B repeat: matches 1. .190 of consensus"
repeat_region   16816. .17125
                /note="AluSg repeat: matches 1. .310 of consensus"
repeat_region   17126. .17338
                /note="MLT1B repeat: matches 190. .381 of consensus"
repeat_region   17506. .17806
                /note="AluYb8 repeat: matches 1. .318 of consensus"
repeat_region   17807. .18030
                /note="7 copies 32 mer 64 conserved"
repeat_region   17808. .18031
                /note="112 copies 2 mer aa 62 conserved"
repeat_region   18046. .18303
                /note="HAL1 repeat: matches 351. .597 of consensus"
repeat_region   18304. .18597
                /note="AluSg repeat: matches 1. .309 of consensus"
repeat_region   18598. .18822
                /note="HAL1 repeat: matches 597. .825 of consensus"
repeat_region   19070. .19305
                /note="L1MC5 repeat: matches 6484. .6730 of consensus"
repeat_region   19306. .19621
                /note="AluSg repeat: matches 1. .310 of consensus"
repeat_region   19622. .20401
                /note="L1MC5 repeat: matches 6730. .7471 of consensus"
repeat_region   20402. .20711
                /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region   20712. .20770
                /note="L1MC5 repeat: matches 7471. .7531 of consensus"
repeat_region   20851. .21158
                /note="AluSg repeat: matches 1. .301 of consensus"
```

```
  Query Match             8.0%;  Score 214.8;  DB 9;  Length 127587;
  Best Local Similarity   80.3%;  Pred. No. 7.5e-46;
  Matches  252;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy       1526 TTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGT 1585
              ||||||||| |||||||||| |||||||| |||||| |||||||||  ||||| | |
Db      59011 TTTGTTTTTCTTTTTTTGAAACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAATGGC 58952

Qy       1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGC 1645
              || ||||| ||| | | || |||| |  || |||  ||||| ||| ||||  |||||||
Db      58951 GCGATCTCGGCTTATTGCAGCCTCTGCCTCCCGGGATTCAAGTGATTCTCCTGCCTCAGC 58892

Qy       1646 CGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTT 1705
              | || |||||||||| | || |||||  ||||| |||||||||||||||||||  |||||||
Db      58891 CTCCCAAGTAGCTGGAATTACAGGTGCATACCACCACACCTGGCTAATTTTTATATTTTT 58832

Qy       1706 TGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAA 1765
               ||||| |  ||||||| |||||||||||||||||||||||| |||||| ||||||  |
Db      58831 AGTAGAGATGGGGTTTCACCATGTTGCCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGA 58772

Qy       1766 CCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGC 1825
              ||||| |||||||||||||||| | |||||||| ||| || |||||||||  | |
Db      58771 TCTGCCTGCCTCGGCCTCCCAAAGTGCCGGGATTACAGGCAAGAGCCACCATGCCGGGCC 58712

Qy       1826 TAGATGTGTCTTAT 1839
              | ||| | | || |
Db      58711 TGGTTTTCTGTTTT 58698
```

```
RESULT 15
AC010352/c
LOCUS       AC010352            102332 bp    DNA     linear   PRI 18-APR-2000
DEFINITION  Homo sapiens chromosome 5 clone CTD-2022F20, complete sequence.
ACCESSION   AC010352
VERSION     AC010352.4  GI:7109394
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 102332)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 102332)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 102332)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 102332)
```

```
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 29, 2000 this sequence version replaced gi:6606056.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 1.5.
            STS Content:
            WI-12799 G24393
            SHGC-13554 G14788.
FEATURES             Location/Qualifiers
     source          1. .102332
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2022F20"
BASE COUNT    29578 a  22934 c  23227 g  26593 t
ORIGIN

  Query Match           8.0%;  Score 214.2;  DB 9;  Length 102332;
  Best Local Similarity   82.6%;  Pred. No. 1e-45;
  Matches  257;  Conservative    0;  Mismatches   53;  Indels    1;  Gaps    1;

Qy       1522 GAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAG 1581
              || ||||  |||| |||||||||| |   ||||||| ||||||||||||||||   |||||
Db      22554 GATTTTTTATTTTTTTTTTTTGAGACACAGTCTCGCTCTGTCGCCCAGGCTGGAATGCAG 22495

Qy       1582 TAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCT 1641
              | ||||||||||| |||| || |||||||||||| ||| ||| | ||||||| ||||  |||
Db      22494 TGGTGCAATCTCGGCTCCCTGCAACCTCCACTTCCT-GGGTTGAAGCGATTCTCCTGCCT 22436

Qy       1642 CAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTAT 1701
              ||||| ||  |||||||||||||| ||| |||| ||| |||||| ||||||||||||||
Db      22435 CAGCCTCCTGAGTAGCTGGGACTACAGGCGTGTGCCACCACACCCGGCTAATTTTTGTAT 22376

Qy       1702 TTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAA 1761
              |||| ||||| |  ||||||| |||||||| |||||||||||| ||| |||||| ||||
Db      22375 TTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAG 22316

Qy       1762 GCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCC 1821
              |||| |||||   |||||||||||||||| |||||||||| |||  || ||||||||||
Db      22315 GCAATCTGCCCATCTCGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCACCATGCCC 22256

Qy       1822 AGGCTAGATGT 1832
              || |||  | |
Db      22255 AGACTATTTAT 22245


Search completed: June 23, 2003, 23:49:50
Job time : 6789 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 20:33:28 ; Search time 568 Seconds
(without alignments)
10601.837 Million cell updates/sec

Title: US-09-817-199B-1
Perfect score: 2674
Sequence: 1 ttcgcctgcgggccggcact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2581.6	96.5	2623	22	AAS27036	cDNA encoding nove
2	2240	83.8	2323	22	AAI60852	Human polynucleoti
3	1286	48.1	1316	22	AAS27458	cDNA encoding nove
4	1099.6	41.1	1106	22	AAI59066	Human polynucleoti
5	1079	40.4	1109	22	AAH75184	Nucleotide sequenc
6	873	32.6	875	24	ABN83696	Human Ras protein
7	795.4	29.7	797	21	AAA52610	Eosinophil activat
8	568	21.2	576	24	ABA02774	Human degranulatio
9	550.8	20.6	843	20	AAX86720	DNA sequence encod
10	480	18.0	576	24	ABA02773	Mouse degranulatio
11	468.6	17.5	475	23	AAS81431	DNA encoding novel
12	292.6	10.9	1340	19	AAV65197	Human RAB protein,
13	274	10.2	443	21	AAA89717	Mouse Exo104 nucle
14	237	8.9	2001	23	ABL29523	Drosophila melanog
15	209.4	7.8	964	22	AAS27040	cDNA encoding nove
16	207.4	7.8	24000	21	AAA88551	Human dual-specifi
17	206.6	7.7	1786	22	AAD08354	Human secreted pro
18	204.8	7.7	10859	22	AAL04720	Human reproductive
19	204.8	7.7	10859	23	ABL97627	Human testicular a
20	203.4	7.6	138169	21	AAA34791	Human adenosine re
21	203.4	7.6	141589	21	AAF20913	Human ELAM-1 polyn
22	203.4	7.6	141589	21	AAF21127	Human low adenosin
23	203.4	7.6	141589	21	AAF21152	Human low adenosin
24	203.4	7.6	141589	21	AAA35005	Human adenosine re
25	203.4	7.6	141589	21	AAA35030	Human adenosine re
26	203.4	7.6	146981	21	AAF21442	Human ELAM-1 polyn
27	203.4	7.6	209273	21	AAF21437	Human factor-relat
c 28	203	7.6	409	22	AAK70203	Human immune/haema
29	202.8	7.6	27579	22	AAK66484	Human immune/haema
c 30	202.6	7.6	29222	24	ABL39412	Human electron-tra
c 31	202.4	7.6	10468	22	AAK77961	Human immune/haema
c 32	201.8	7.5	103747	24	ABQ88139	Human osteoblast d
33	201.6	7.5	52845	22	AAK71437	Human immune/haema
34	201.4	7.5	4969	22	ABA16169	Human nervous syst
35	201.4	7.5	6720	24	ABQ88138	Human osteoblast d
c 36	201.2	7.5	33353	22	AAK70003	Human immune/haema
37	201	7.5	7428	22	ABA07164	Human pancreatic c
38	200.8	7.5	21835	22	AAK66483	Human immune/haema
c 39	200.6	7.5	21833	22	AAL35910	Human musculoskele
40	200.4	7.5	10663	22	ABA18961	Human nervous syst
c 41	200.4	7.5	31865	22	AAK85330	Human immune/haema
c 42	200.4	7.5	31865	22	AAK85331	Human immune/haema
c 43	200.2	7.5	21777	22	AAK74096	Human immune/haema
44	199.8	7.5	1130	22	ABA20629	Human nervous syst
c 45	199.8	7.5	1130	22	AAL37008	Human musculoskele

ALIGNMENTS

RESULT 1
AAS27036
ID AAS27036 standard; cDNA; 2623 BP.
XX
AC AAS27036;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 71.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR P-PSDB; AAU17119.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 71; 880pp; English.

XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Query Match 96.5%; Score 2581.6; DB 22; Length 2623;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy        42 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 71

Qy       102 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        72 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 131

Qy       162 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       132 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 191

Qy       222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 281
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       192 GCCACCGTCGGCATAGACTTCAGCAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 251

Qy       282 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       252 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 311

Qy       342 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       312 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 371

Qy       402 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       372 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 431

Qy       462 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       432 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 491

Qy       522 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       492 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 551

Qy       582 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       552 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 611

Qy       642 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       612 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 671

Qy       702 TCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCT 761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       672 TCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCT 731

Qy       762 TCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGA 821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       732 TCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGA 791

Qy       822 CTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGC 881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       792 CTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGC 851

Qy       882 GGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAAT 941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       852 GGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAAT 911

Qy       942 TTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTT 1001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       912 TTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTT 971

Qy      1002 TTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGC 1061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       972 TTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGC 1031

Qy      1062 TCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCA 1121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1032 TCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCA 1091

Qy      1122 TCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGC 1181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1092 TCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGC 1151

Qy      1182 TCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAA 1241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1152 TCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAA 1211

Qy      1242 GACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGA 1301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1212 GACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGA 1271

Qy      1302 GAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACA-CCCCCCTGGAGGCATGCCCCT 1360
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      1272 GAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCCTGGAGGCATGCCCCT 1331

Qy      1361 TTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCT 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1332 TTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCT 1391

Qy      1421 CTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTC 1480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1392 CTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTC 1451

Qy      1481 CTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTT 1540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1452 CTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTT 1511

Qy      1541 TTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCAC 1600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1512 TTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCAC 1571

Qy      1601 TACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGG 1660
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      1572 TACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCCCCGAAGTAGCTGG 1631

Qy      1661 GACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTT 1720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1632 GACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTT 1691

Qy      1721 TCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGC 1780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1692 TCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGC 1751

Qy      1781 CTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATC 1840
```

```
Db       1752 CTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATC 1811

Qy       1841 CCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTA 1900

Db       1812 CCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTA 1871

Qy       1901 GCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTG 1960

Db       1872 GCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTG 1931

Qy       1961 TG-AAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 2019

Db       1932 TGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 1991

Qy       2020 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 2079

Db       1992 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 2051

Qy       2080 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 2139

Db       2052 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 2111

Qy       2140 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 2199

Db       2112 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 2171

Qy       2200 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 2259

Db       2172 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 2231

Qy       2260 TGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 2319

Db       2232 TGGTCACATATCTCCCAAAGCTCCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 2291

Qy       2320 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGA 2379

Db       2292 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGAGGAGAGGGGGAGGA 2351

Qy       2380 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 2439

Db       2352 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 2411

Qy       2440 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 2499

Db       2412 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 2471

Qy       2500 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 2559

Db       2472 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 2531

Qy       2560 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 2619

Db       2532 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 2591

Qy       2620 GCAGTATTTTGGAAAGCAAAAAAAAAAAAAAA 2651

Db       2592 GCAGTATTTTGGAAAGCAAAAAAAAAAAAAA 2623
```

RESULT 2
AAI60852
ID AAI60852 standard; cDNA; 2323 BP.
XX
AC AAI60852;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4841.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41696.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4841; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Query Match 83.8%; Score 2240; DB 22; Length 2323;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2298; Conservative 0; Mismatches 16; Indels 9; Gaps 4;

```
Qy         15 GGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 74

Db          1 GGAATTCCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 60

Qy         75 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 134

Db         61 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 120

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194

Db        121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 180

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254

Db        181 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240
```

```
Qy       255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
Db       241 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 300

Qy       315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
Db       301 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 360

Qy       375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
Db       361 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 420

Qy       435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 494
Db       421 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 480

Qy       495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 554
Db       481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 540

Qy       555 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 614
Db       541 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 600

Qy       615 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 674
Db       601 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 660

Qy       675 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 734
Db       661 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 720

Qy       735 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 794
Db       721 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 780

Qy       795 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 854
Db       781 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 840

Qy       855 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 914
Db       841 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 900

Qy       915 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 974
Db       901 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 960

Qy       975 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1034
Db       961 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1020

Qy      1035 TCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1094
Db      1021 TCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1080

Qy      1095 TGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAG 1154
Db      1081 TGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAG 1140

Qy      1155 GTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAA 1214
Db      1141 GTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAA 1200

Qy      1215 ATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGA 1274
Db      1201 ATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGA 1260

Qy      1275 GATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTAC 1334
Db      1261 GATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTAC 1320

Qy      1335 TAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGG 1394
Db      1321 TAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGG 1380

Qy      1395 AAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCC 1454
Db      1381 AAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCC 1440

Qy      1455 CCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACT 1514
Db      1441 CCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACT 1500

Qy      1515 AGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGA 1574
Db      1501 AGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGA 1560

Qy      1575 GGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1634
Db      1561 GGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1620

Qy      1635 CCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATT 1694
Db      1621 CCCACCTCAGCCCCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATT 1680

Qy      1695 TTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTG 1754
Db      1681 TTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTG 1740

Qy      1755 AGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCAC 1814
Db      1741 AGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCAC 1800

Qy      1815 CATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCG 1874
Db      1801 CATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCG 1860

Qy      1875 ATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAAT 1934
Db      1861 ATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAAT 1920

Qy      1935 CCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCT 1994
Db      1921 CCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCT 1980

Qy      1995 GGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2054
Db      1981 GGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2040

Qy      2055 CCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAA 2114
Db      2041 CCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAA 2100

Qy      2115 ACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGG---TCACCCATTTC 2171
Db      2101 ACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGGAGGTTCAACCCATTTC 2160

Qy      2172 CGAGTT--AAACCAATGCAATA-TGAGTAAAACAAAGT---CATGTGGGTATGTCTGGGG 2225
Db      2161 CGACTTTAAACCCAATGCAATATTGAGTAAAACAAAGTTCATGTGGGGTATTGTCTGGGG 2220

Qy      2226 TAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGA 2285
Db      2221 TAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGA 2280

Qy      2286 TCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCT 2328
Db      2281 TCCCTNCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCT 2323


RESULT 3
AAS27458
ID   AAS27458 standard; cDNA; 1316 BP.
XX
```

AC AAS27458;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 493.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

```
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465460/50.
DR   P-PSDB; AAU17541.
XX
PT   Novel polypeptides useful for diagnosing, treating, preventing and/or
PT   prognosing disorders related to the proteins, including cancers, immune
PT   disorders and neuronal disorders  -
XX
PS   Claim 1; SEQ ID No 493; 880pp; English.
XX
CC   The invention relates to novel isolated polypeptides (I), and
CC   polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC   diagnosing, preventing and treating diseases including immune system
CC   disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC   disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC   transplant rejections and graft versus host disease, infectious diseases
CC   (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC   other blood-related disorders (sickle cell anaemia), myeloproliferative
CC   disorders, primary haematopoietic disorders, hyperproliferative
CC   disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC   disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC   abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC   disorders (e.g. glomerulonephritis), cardiovascular disorders
CC   (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC   wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC   Addison's disease), reproductive system disorders, gastrointestinal
CC   disorder (inflammatory disorders), liver disorders (cirrhosis),
CC   as stimulators of B-cell responsiveness to pathogens, activators of
CC   T-cells, to induce higher affinity antibodies, and as a means to induce
CC   tumour proliferation in pathologies e.g. acquired immune deficiency
CC   syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC   pathway protein coding sequences and PCR primers of the invention.
XX

  Query Match             48.1%;  Score 1286;  DB 22;  Length 1316;
  Best Local Similarity   98.5%;  Pred. No. 5.6e-237;
  Matches 1286;  Conservative    9;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         45 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 60

Qy        105 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 120

Qy        165 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 180

Qy        225 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 284
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        181 ACCGTCGGCATAGACTTCAGSAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 240

Qy        285 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 300

Qy        345 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 360

Qy        405 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 420

Qy        465 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 480

Qy        525 GCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 584
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db        481 GCCAGGGAGTACGGKGTTCCCTTYCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 540

Qy        585 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 600

Qy        645 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 704
              ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        601 CCCAGCTTCCAGATCCGAGACTATGTAKAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 660

Qy        705 TTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCC 764
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        661 TTCATGTGAATCCCAGGGGGCAKAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCC 720

Qy        765 CCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTC 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTC 780

Qy        825 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGT 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGT 840

Qy        885 CTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTA 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTA 900

Qy        945 ATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTG 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTG 960

Qy       1005 CTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCC 1064
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTTCTAGGACTTGGGSGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCC 1020

Qy       1065 AGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCG 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCG 1080

Qy       1125 CACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCT 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCT 1140

Qy       1185 CCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGAC 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGAC 1200

Qy       1245 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAG 1260

Qy       1305 TAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAG 1350
              ||||||:|||||||:|||||| ||| |||: | |   || ||  | |
Db        1261 TAGGGTYTAGCTGGYTATCTNTGGGCTTWATTAACACCNCCCTGGG 1306
```

```
RESULT 4
AAI59066
ID   AAI59066 standard; cDNA; 1106 BP.
XX
AC   AAI59066;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1269.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39910.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 1269; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 1106 BP; 260 A; 307 C; 321 G; 218 T; 0 other;

  Query Match             41.1%;  Score 1099.6;  DB 22;  Length 1106;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-201;
  Matches 1102;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 GGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 74
              || | | ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 GGAAGTCCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 60

Qy         75 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 120

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 180

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240

Qy        255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 300

Qy        315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        301 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 360

Qy        375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 420

Qy        435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 480

Qy        495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 540

Qy        555 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        541 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 600

Qy        615 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        601 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 660

Qy        675 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        661 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 720

Qy        735 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        721 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 780

Qy        795 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        781 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 840

Qy        855 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        841 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 900

Qy        915 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        901 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 960

Qy        975 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        961 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1020

Qy       1035 TCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1094
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||

Db       1021 TCCTAAGCATAACAAAGGTGGGGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1080

Qy       1095 TGGGGGTGAGGGCAGGTAATGACTCC 1120
              |||||||||||||||||||||||||
Db       1081 TGGGGGTGAGGGCAGGTAATGACTCC 1106
```

```
RESULT 5
AAH75184
ID   AAH75184 standard; DNA; 1109 BP.
XX
AC   AAH75184;
XX
DT   13-NOV-2001  (first entry)
XX
DE   Nucleotide sequence of human 32712 G-protein.
XX
KW   Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW   congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW   adult respiratory distress syndrome; Goodpasture's syndrome;
KW   chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
KW   liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
KW   cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
KW   brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
KW   acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
KW   chronic bacterial meningoencephalitis; multiple sclerosis;
KW   amyotropic lateral sclerosis; stroke; Huntington's disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             124..699
FT                   /*tag= a
FT                   /product= "G-protein"
XX
PN   WO200164887-A2.
XX
PD   07-SEP-2001.
XX
PF   27-FEB-2001; 2001WO-US06292.
XX
PR   29-FEB-2000; 2000US-0185606.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers RA;
XX
DR   WPI; 2001-550182/61.
DR   P-PSDB; AAG67156.
XX
PT   Novel human small G-protein polypeptides and polynucleotides for
PT   treating lung disorders, liver disorders and brain disorders -
XX
PS   Claim 2; Fig 26; 151pp; English.
XX
CC   The present sequence encodes a human G-protein. The specification
CC   describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
CC   G-protein polypeptides and polynucleotides are useful as a target for
CC   diagnosis and treatment of  G-protein mediated or related disorders,
CC   and for identifying agonists and antagonists for diagnosis and
CC   treatment. They are useful for treating disorders of lung (e.g.
CC   congenital anomalies, pulmonary congestion, oedema, adult respiratory
CC   distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
CC   asthma, Goodpasture's syndrome and pulmonary hypertension), liver
CC   (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
CC   Wilson's disease, autoimmune hepatitis and hepatic failure), and
CC   brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
CC   meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
CC   bacterial meningoencephalitis, multiple sclerosis, amyotropic lateral
CC   sclerosis, stroke and Huntington's disease).
XX
SQ   Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

  Query Match            40.4%;  Score 1079;  DB 22;  Length 1109;
  Best Local Similarity  99.5%;  Pred. No. 2.2e-197;
  Matches 1082;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         37 GGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 96
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 GCGTCCGGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 82

Qy         97 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 156
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 142

Qy        157 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 216
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 202

Qy        217 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 276
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 262

Qy        277 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 336
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 322

Qy        337 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 396
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 382

Qy        397 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 456
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 442

Qy        457 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 516
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 502

Qy        517 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 576
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 562

Qy        577 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 636
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 622

Qy        637 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 696
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 682

Qy        697 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 756
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 742

Qy        757 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 816
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 802

Qy        817 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 876
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 862

Qy        877 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 936
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 922

Qy        937 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 996
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 982

Qy        997 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1056
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1042

Qy       1057 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGA 1116
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1043 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAAGTAATGA 1102

Qy       1117 CTCCATC 1123
                |||||||
```

```
Db      1103 CTCCATC 1109


RESULT 6
ABN83696
ID   ABN83696 standard; cDNA; 875 BP.
XX
AC   ABN83696;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human Ras protein 3 (RASP-3) cDNA.
XX
KW   Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;
KW   immunosuppressive; antiinflammatory; signal transduction; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             93..668
FT                   /*tag= a
FT                   /product= "RASP-3"
XX
PN   US6391580-B1.
XX
PD   21-MAY-2002.
XX
PF   08-MAY-1998;   98US-0075454.
XX
PR   12-DEC-1996;   96US-0766551.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Tang YT,  Lal P,  Guegler KJ,  Corley NC,  Patterson C;
PI   Batra S,  Baughn MR;
XX
DR   WPI; 2002-498774/53.
DR   P-PSDB; ABB76426.
XX
PT   New isolated polynucleotides encoding Ras proteins designated RASP-1
PT   and RASP-4, for diagnosing, preventing and treating disorders
PT   associated with cell proliferation, particularly cancer and immune
PT   disorders  -
XX
PS   Example; Column 57-60; 34pp; English.
XX
CC   The present sequence is that of a cDNA clone encoding novel human
CC   Ras protein 3 (RASP-3, see ABB76426).  Nucleic acids encoding
CC   RASP-3 were initially identified in Incyte Clone 1528559 from a
CC   mononuclear cell cDNA library (UCMCL5T01) using a computer search
CC   for amino acid sequence alignments.  The present consensus
CC   sequence was derived from overlapping and/or extended nucleic
CC   acids in Incyte clone 1528559 and shot-gun sequence SAEA03135,
CC   SAEC10396 and SAEC10855.  RASP-3 shows homology to rat Rab26.
CC   Northern analysis showed expression of RASP-3 in haematopoietic
CC   and immunological cDNA libraries, all associated with inflammation
CC   and the immune response.  The invention provides 7 novel human Ras
CC   proteins (RASP-1 to -7) and polynucleotides, expression vectors,
CC   host cells, antibodies, agonists and antagonists.  It also provides
CC   methods for diagnosing, treating or preventing disorders associated
CC   with RASP expression, especially cancer and immune disorders.  A
CC   fragment of the present sequence, from about nucleotide 92 to about
CC   nucleotide 153, is useful as a hybridisation probe.
XX
SQ   Sequence 875 BP; 203 A; 247 C; 260 G; 165 T; 0 other;

  Query Match            32.6%;  Score 873;  DB 24;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-158;
  Matches  873;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 60

Qy        106 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 120

Qy        166 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 180

Qy        226 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 240

Qy        286 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 300

Qy        346 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        301 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 360

Qy        406 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 420

Qy        466 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 480

Qy        526 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        481 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 540

Qy        586 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        541 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 600

Qy        646 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        601 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 660

Qy        706 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        661 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 720

Qy        766 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        721 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 780

Qy        826 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        781 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 840

Qy        886 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 918
              ||||||||||||||||||||||||||||||||
Db        841 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 873


RESULT 7
AAA52610
ID   AAA52610 standard; DNA; 797 BP.
XX
AC   AAA52610;
XX
DT   07-DEC-2000  (first entry)
XX
DE   Eosinophil activating peptide gene 8.
XX
KW   Eosinophil activation; human; allergy; eosinophilia; cancer;
KW   inflammation; signalling peptide; cell adhesion peptide;
KW   G-protein coupled receptor; secreted protein; cell surface protein; ds.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200032630-A2.
XX
PD   08-JUN-2000.
XX
PF   06-DEC-1999;   99WO-US28773.
XX
PR   04-DEC-1998;   98US-0111006.
XX
PA   (SEAR ) SEARLE & CO G D.
XX
PI   Dotson SB,  Ma X;
XX
DR   WPI; 2000-465041/40.
XX
PT   Novel nucleic acids derived from activated eosinophil cells useful for
PT   treating allergic diseases such as asthma comprises a specific
PT   nucleotide sequence  -
XX
PS   Claim 1; Page 95; 125pp; English.
XX
CC   The present invention relates to a number of nucleotide sequences
CC   which encode proteins involved in the activation of eosinophils
CC   (AAA52601-A52679). Eosinophils are involved in immune reactions, and
CC   these genes and their proteins provide possible targets for new drugs to
CC   combat diseases such as asthma, allergic rhinitis, atopic dermatitis,
CC   anaphylaxis, allergic bronchopulmonary aspergillosis, eczema, psoriasis,
CC   emphysema, leukaemia, lymphomas, ovarian cancer, pneumonia, immune
CC   disorders, some connective tissue disorders, and inflammatory conditions
CC   including septic shock, arthritis, nephritis, inflammatory bowel disease
CC   and Crohn's disease.
XX
SQ   Sequence 797 BP; 192 A; 201 C; 209 G; 195 T; 0 other;

  Query Match            29.7%;  Score 795.4;  DB 21;  Length 797;
  Best Local Similarity  99.9%;  Pred. No. 4e-143;
  Matches  796;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1840 CCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTT 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db          1 CCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGCAGTGTTT 60

Qy       1900 AGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTT 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTT 120

Qy       1960 GTGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 2019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 180

Qy       2020 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 2079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 240

Qy       2080 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 2139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 300

Qy       2140 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 2199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 360

Qy       2200 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 2259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 420

Qy       2260 TGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 2319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 480

Qy       2320 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGA 2379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGA 540

Qy       2380 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 2439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 600

Qy       2440 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 2499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 660

Qy       2500 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 2559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 720

Qy       2560 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 2619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 780

Qy       2620 GCAGTATTTTGGAAAGC 2636
             |||||||||||||||||
Db        781 GCAGTATTTTGGAAAGC 797


RESULT 8
ABA02774
ID   ABA02774 standard; DNA; 576 BP.
XX
AC   ABA02774;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human degranulation regulator encoding DNA SEQ ID NO 3.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..576
FT                   /*tag=  a
FT                   /product= "degranulation regulator"
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
DR   P-PSDB; AAM52190.
XX
PT   Mast cell degranulation controller for treatment of allergies  -
XX
PS   Claim 11; Page 64-65; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;

  Query Match            21.2%;  Score 568;  DB 24;  Length 576;
  Best Local Similarity  99.1%;  Pred. No. 1.2e-99;
  Matches  571;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 197
```

```
Db          1 ATGCTTCTTGGAGACTCGGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60

Qy        198 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy        258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy        318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        378 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        438 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        498 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        558 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        618 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        678 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 713
              ||||||||||||||||||||||||||||||  |||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTGTGA 576


RESULT 9
AAX86720
ID   AAX86720 standard; DNA; 843 BP.
XX
AC   AAX86720;
XX
DT   27-OCT-1999  (first entry)
XX
DE   DNA sequence encoding an exocytotic protein designated Exo2.
XX
KW   Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
KW   inflammation; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             124..699
FT                   /*tag= a
FT                   /transl_except= (pos: 136..138, aa: Xaa)
FT                   /transl_except= (pos: 202..204, aa: Xaa)
FT                   /note= "Xaa is an unspecified amino acid"
XX
PN   WO9942586-A2.
XX
PD   26-AUG-1999.
XX
PF   23-FEB-1999;   99WO-US03909.
XX
PR   26-MAY-1998;   98US-0086650.
PR   23-FEB-1998;   98US-0075534.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Fisher J,  Huang B,  Lorens J,  Luo Y;
PI   Shen M;
XX
DR   WPI; 1999-518605/43.
DR   P-PSDB; AAY30132.
XX
PT   New exocytotic proteins useful for diagnosis and treatment of
PT   exocytosis-mediated conditions and in drug screening
XX
PS   Claim 6; Fig 6; 53pp; English.
XX
CC   The present sequence encodes an exocytotic protein designated Exo2.
CC   The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC   secretory vesicles with the cellular plasma membrane) and can be
CC   administered therapeutically to treat or prevent exocytosis-mediated
CC   disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC   undesirable release of compounds via exocytosis e.g. inflammation
CC   mediated by the release compounds such as histamine. Exo2 is useful
CC   diagnostically and to produce antibodies useful to purify the proteins
CC   and therapeutically to reduce or eliminate the biological activity of
CC   the protein. Exo2 nucleic acids can be used therapeutically to increase
CC   Exo2 activity in cells by known gene therapy techniques. They can also
CC   be used to produce probes or primers to isolate Exo2 proteins from other
CC   organisms, especially humans. The nucleic acids, host cells and proteins
CC   are useful in screening assays to identify binding agents, especially
CC   drug screening assays to identify agonists and antagonists useful
CC   therapeutically to enhance or reduce Exo2 activity.
XX
SQ   Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

  Query Match             20.6%;  Score 550.8;  DB 20;  Length 843;
  Best Local Similarity   86.8%;  Pred. No. 2.5e-96;
  Matches  612;  Conservative    0;  Mismatches   92;  Indels    1;  Gaps    1;

Qy         73 GGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCA 132
              |||||||||||||||| |||||||||||||||| ||| |||| ||||||  |||| ||||
Db         59 GGGATGGCGAGGCCCCTGAGCGCTCCCCGCCCTTCAGCCCGAACTACGATNTCACCGGCA 118

Qy        133 AGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAG 192
              |||||||||||||| |||||  | |||||||||||||| ||||||||||||||||||||
Db        119 AGGTGATGCTTCTTGGAGANTCGGGCGTCGGCAAAACCTGTTTCCTGATCCAATTCAAAG 178

Qy        193 ACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGG 252
              |||||||||||||||||||||||| |||||||||||||||||||||||||||| || |
Db        179 ACGGGGCCTTCCTGTCCGGAACCTTNATAGCCACCGTCGGCATAGACTTCAGGAATAAAG 238

Qy        253 TGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAAC 312
              ||||||| ||||||||    || |||||||| |||||||||||||| || || ||||| |
Db        239 TGGTGACAGTGGATGGTTCCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGGAGC 298

Qy        313 GGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATG 372
              | ||||| || || |||||||||||||| |||||||||||| ||||| ||| ||||||
Db        299 GCTTCCGCAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCCTGTTGTATG 358

Qy        373 ACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGT 432
              |||||||||| | || ||||| |||||||||||||||||||||| ||||||||||||
Db        359 ACATCACCAACCAGTCCTCTTTTGACAACATCAGGGCCTGGCTCACAGAGATTCATGAGT 418

Qy        433 ATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAA 492
              |||||||||||||||||||| ||||| |||||||||||||| ||| | ||||||||||
Db        419 ATGCCCAGAGGGACGTGGTGATTATGCTTCTAGGCAACAAGGCCGATGTAAGCAGCGAAA 478

Qy        493 GAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGG 552
              | |||||||||| ||||| ||||||||  |||||||||| || |||||||| ||| |||
Db        479 GGGTGATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAATATGGTGTTCCTTTCATGG 538

Qy        553 AGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAAC 612
              ||||||| |||||||||||||||| |||||||| |||||||||| || ||||||||||
Db        539 AGACCAGTGCCAAGACTGGCATGAACGTGGAGTTGGCCTTTCTGGCAATTGCCAAGGAAC 598

Qy        613 TGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAG 672
              |||||||||| || |||   ||| | |||||||||||||||||||||||||||||| |
Db        599 TGAAATACCGTGCAGGGAGGCAGCCTGATGAGCCCAGCTTCCAGATCCGAGACTATGTGG 658
```

```
Qy        673 AGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGG 732
              ||||||||||||||||||||||||||||||||||  |||||  |||  |||||  |||||
Db        659 AGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTGTGACCCCCTAGGGGCTAAGAGG 718

Qy        733 AGGCTCTGGAGGCAC-ACAGGATGCAGCCTTCCCCCTCCCAGGCC 776
              |||  |   || | |    ||||||||   |||  || |||  ||
Db        719 AGGGCCCANAGACCCTTGGGGATGCAGTACTCCAACTGCCACACC 763


RESULT 10
ABA02773
ID   ABA02773 standard; DNA; 576 BP.
XX
AC   ABA02773;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Mouse degranulation regulator encoding DNA SEQ ID NO 2.
XX
KW   Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..576
FT                   /*tag= a
FT                   /transl_except= (pos:247..249,aa:Lys)
FT                   /product= "degranulation regulator"
XX
PN   WO200179478-A1.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-JP03268.
XX
PR   19-APR-2000; 2000JP-0118408.
XX
PA   (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI   Yamada T,  Ido M;
XX
DR   WPI; 2002-041335/05.
DR   P-PSDB; AAM52189.
XX
PT   Mast cell degranulation controller for treatment of allergies -
XX
PS   Claim 10; Page 63-64; 85pp; Japanese.
XX
CC   The invention relates to a protein for regulating degranulation of mast
CC   cells (degranulation regulators) and the encoding polynucleotides, with
CC   antiallergic activity, used in the treatment of allergies associated
CC   with degranulation of mast cells.
XX
SQ   Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

  Query Match           18.0%;  Score 480;  DB 24;  Length 576;
  Best Local Similarity   89.6%;  Pred. No. 8.6e-83;
  Matches  516;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy        138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 197
              |||||||| |||||| | ||||||||||||||| ||||||||||||||||||||||||||
Db          1 ATGCTTCTTGGAGACTCGGGCGTCGGCAAAACCTGTTTCCTGATCCAATTCAAAGACGGG 60

Qy        198 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTG 120

Qy        258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 317
              || |||||||| |  || |||||||| |||||||||||||| || || ||||| || ||
Db        121 ACAGTGGATGGTGCCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGGAGCGCTTT 180

Qy        318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 377
              || || || |||||||||||||| |||||||||||| ||||| ||| ||||||||||
Db        181 CGCAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCCTGTTGTATGACATC 240

Qy        378 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 437
              |||||| | || ||||| |||||||||||||||||||||| ||||||||||||||||||
Db        241 ACCAACCAGTCCTCTTTTGACAACATCAGGGCCTGGCTCACAGAGATTCATGAGTATGCC 300

Qy        438 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 497
              ||||| |||||||||| ||||| |||||||||||||| ||| | |||||||||| |||
Db        301 CAGAGAGACGTGGTGATTATGCTTCTAGGCAACAAGGCCGATGTAAGCAGCGAAAGGGTG 360

Qy        498 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 557
              |||||||| ||||| |||||||| |||||||||| || |||||||| ||| ||||||||
Db        361 ATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAATATGGTGTTCCTTTCATGGAGACC 420

Qy        558 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 617
              || |||||||||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        421 AGTGCCAAGACTGGCATGAACGTGGAGTTGGCCTTTCTGGCAATTGCCAAGGAACTGAAA 480

Qy        618 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
              ||||| || |||   ||| | ||||||||||||||||||||||||||||||| ||||||
Db        481 TACCGTGCAGGGAGGCAGCCTGATGAGCCCAGCTTCCAGATCCGAGACTATGTGGAGTCC 540

Qy        678 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 713
              |||||||||||||||||||||||||||||  |||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTGTGA 576


RESULT 11
AAS81431
ID   AAS81431 standard; cDNA; 475 BP.
XX
AC   AAS81431;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #17235.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17244.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 17235; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
```

```
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;

  Query Match            17.5%;  Score 468.6;  DB 23;  Length 475;
  Best Local Similarity  99.2%;  Pred. No. 1.3e-80;
  Matches  471;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 GGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 74
              || | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAATTCCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 60

Qy         75 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 120

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 180

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240

Qy        255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 300

Qy        315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 360

Qy        375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 420

Qy        435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCG 489
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCCAGAAGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCG 475


RESULT 12
AAV65197
ID   AAV65197 standard; DNA; 1340 BP.
XX
AC   AAV65197;
XX
DT   17-DEC-1998  (first entry)
XX
DE   Human RAB protein, SRAB, coding sequence.
XX
KW   Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW   Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW   viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW   diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW   hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;
KW   polycystic renal disease; ss.
XX
OS   Homo sapiens.
XX

FH   Key             Location/Qualifiers
FT   CDS             43..615
FT                   /*tag= a
FT                   /product= "SRAB"
XX
PN   WO9842839-A1.
XX
PD   01-OCT-1998.
XX
PF   25-MAR-1998;    98WO-US05871.
XX
PR   26-MAR-1997;    97US-0824873. .
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Guegler KJ,  Hillman JL;
XX
DR   WPI; 1998-542282/46.
DR   P-PSDB; AAW80747.
XX
PT   New human Rab protein, SRAB - useful for treating disorders
PT   associated with SRAB expression, including vesicle trafficking,
PT   viral infection, and cancer
XX
PS   Claim 5; Fig 1; 31pp; English.
XX
CC   This sequence encodes the human RAB protein, SRAB, of the invention. The
CC   SRAB protein can be used in compositions useful in the diagnosis,
CC   prevention, or treatment of disorders associated with vesicle trafficking
CC   (including epilepsy, Huntington's disease, Parkinson's disease and
CC   schizophrenia), cancer, or viral infection. Host cells containing the DNA
CC   sequence can be used to produce SRAB recombinantly. The viral infections
CC   that can be treated using SRAB include those caused by retroviruses
CC   particularly HIV and HTLV, hepadnaviruses, particularly hepatitis C,
CC   hantaviruses, herpesviruses, and arboviruses. The cancers that can be
CC   treated include adenocarcinoma, leukaemia, lymphoma, melanoma and
CC   sarcoma, particularly cancers of the endocrine, gastrointestinal and
CC   nervous systems and cancers of the adrenal gland, brain, breast, colon,
CC   oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,
CC   prostate, salivary gland, stomach, thyroid, and uterus. The protein can
CC   also be used to treat conditions such as autoimmune sialosis,
CC   choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,
CC   hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme
CC   deficiency, pancreatitis, and polycystic renal disease.
XX
SQ   Sequence 1340 BP; 276 A; 387 C; 403 G; 274 T; 0 other;

  Query Match            10.9%;  Score 292.6;  DB 19;  Length 1340;
  Best Local Similarity  70.9%;  Pred. No. 6.8e-47;
  Matches  416;  Conservative    0;  Mismatches  169;  Indels    2;  Gaps    2;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              ||||||| || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db         21 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 80

Qy        176 CCTGA-TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCA 234
               |||  | | |||||| || || || |||||| | || ||||||||  |||||||  |||
Db         81 GCTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139

Qy        235 TAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGG 294
              | |||||| ||||||| ||  ||   |||||||| ||||  ||||||||||||||| ||||
Db        140 TTGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGG 199

Qy        295 ACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGG 354
              |||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| |
Db        200 ACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATG 259

Qy        355 CCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGC 414
              |  |||| ||||| || ||  |||||||||  | || || |||||||||  ||||||||
Db        260 CTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGC 319

Qy        415 TCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGG 474
              | || ||||| || ||||| ||||||   ||||||| | ||||||||||| || |||||||
```

```
Db        320 TGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGG 379

Qy        475 CGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGT 534
              |||       |   ||  | ||| |        || ||||| ||||   |||||| |||||
Db        380 TGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGT 439

Qy        535 ACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTC 594
              | ||  | |||||| ||||||||||||||||||| ||| | || ||||| || |||||
Db        440 ATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCA 499

Qy        595 TGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCC 654
                |||||| || |||||  ||||  | ||  ||  | |     |    |||||  ||||||
Db        500 CAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCC 559

Qy        655 AGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              | | |  || || ||  ||    ||        |  ||  |||||||
Db        560 GGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606


RESULT 13
AAA89717
ID   AAA89717 standard; cDNA; 443 BP.
XX
AC   AAA89717;
XX
DT   08-JAN-2001  (first entry)
XX
DE   Mouse Exo104 nucleotide sequence.
XX
KW   Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;
KW   antiallergic; antiasthmatic; nootropic; neuroprotective; anticonvulsant;
KW   vulnerary; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
KW   digestion disorder; wound healing disorder; gene therapy; ss.
XX
OS   Mus sp.
XX
PN   WO200043419-A2.
XX
PD   27-JUL-2000.
XX
PF   20-JAN-2000; 2000WO-US01431.
XX
PR   20-JAN-1999;   99US-0116534.
PR   26-JAN-1999;   99US-0117274.
PR   26-JAN-1999;   99US-0117308.
PR   26-JAN-1999;   99US-0117309.
PR   26-JAN-1999;   99US-0117312.
PR   01-FEB-1999;   99US-0118177.
PR   01-FEB-1999;   99US-0118178.
PR   01-FEB-1999;   99US-0118179.
PR   09-FEB-1999;   99US-0119286.
PR   11-FEB-1999;   99US-0119998.
PR   11-FEB-1999;   99US-0119759.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y;
XX
DR   WPI; 2000-482908/42.
XX
PT   New nucleic acids encoding Exo proteins which are useful in the
PT   diagnosis, treatment or prevention of exocytosis-mediated disorders
PT   such as asthma, inflammation and allergies -
XX
PS   Disclosure; Page 290; 305pp; English.
XX
CC   The present sequence encodes a polypeptide which is associated with
CC   the exocytosis pathway. cDNA molecules encoding proteins involved in
CC   exocytosis have been isolated by yeast one-hybrid and two-hybrid
CC   screening. Novel proteins, termed Exo proteins, have been identified that
CC   interact with known exocytosis-associated proteins such as GS27, alpha
CC   snap, unc18-1, vamp3, snap-23, and the rab family of proteins.
CC   Exo proteins and their agonists and antagonists are useful in the
CC   diagnosis, treatment or prevention of exocytosis-mediated disorders
CC   such as asthma, inflammation, allergies, Chediak-Higashi Syndrome
CC   (CHS), Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC   diabetes, digestion disorders and wound healing disorders.
CC   The nucleic acids, antagonists or agonists of Exo proteins are useful
CC   in gene therapy. The nucleic acids are also useful for generating
CC   transgenic or knock-out animals which can be used in the
CC   development and screening of therapeutically useful reagents.
XX
SQ   Sequence 443 BP; 109 A; 111 C; 131 G; 90 T; 2 other;

  Query Match            10.2%;  Score 274;  DB 21;  Length 443;
  Best Local Similarity   84.6%;  Pred. No. 2.2e-43;
  Matches  307;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy        414 CTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAG 473
              ||||| |||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db          1 CTCACAGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATTATGCTTCTAGGCAACAAG 60

Qy        474 GCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAG 533
              || ||| | |||||||||||| |||||||||| ||||| ||||||||  ||||||||||
Db         61 GCCGATGTAAGCAGCGAAAGGGTGATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAA 120

Qy        534 TACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTT 593
              || |||||||| ||| |||||||||| |||||||||||||||| |||||||| ||||||
Db        121 TATGGTGTTCCTTTCATGGAGACCAGTGCCAAGACTGGCATGAACGTGGAGTTGGCCTTT 180

Qy        594 CTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTC 653
              ||||| || |||||||||||||||||||| || |||   ||| | ||||||||||||||||
Db        181 CTGGCAATTGCCAAGGAACTGAAATACCGTGCAGGGAGGCAGCCTGATGAGCCCAGCTTC 240

Qy        654 CAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 713
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||||
Db        241 CAGATCCGAGACTATGTGGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTTGTGTGA 300

Qy        714 ATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAG 773
                |||  |||||  ||||||||| |     |    ||||||||   |||  || |||
Db        301 CCCCCTAGGGGCTAAGAGGAGGCCCANAGACCCTTGGGGATGCAGTACTCCAACTGCCAC 360

Qy        774 GCC 776
                ||
Db        361 ACC 363


RESULT 14
ABL29523
ID   ABL29523 standard; DNA; 2001 BP.
XX
AC   ABL29523;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 40042.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
```

```
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 40042; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2001 BP; 455 A; 616 C; 573 G; 357 T; 0 other;

  Query Match             8.9%;  Score 237;  DB 23;  Length 2001;
  Best Local Similarity   65.4%;  Pred. No. 3.1e-36;
  Matches  364;  Conservative    0;  Mismatches  190;  Indels    3;  Gaps    1;

Qy        117 TACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTC 176
              | |||  ||| ||||||||| ||  | ||||| ||  | || || ||||| || |   |
Db       1393 TTCGATATCATGGGCAAGGTCATAATGCTGGGCGATTCGGGAGTGGGCAAGACCTCACTT 1452

Qy        177 CTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATA 236
              ||||||  |||   || ||   ||   || |  |     ||| |   ||| || |||||
Db       1453 TTGATCCGTTTCCGCGATGGTCGCTATGTGCCGAGCTATTTCCTCAGCACAGTTGGCATT 1512

Qy        237 GACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGAC 296
              || || |||||||| ||||||   || |||||   | | || |||||||| |||||||||
Db       1513 GATTTTAGGAACAAAGTGGTGGTCGTCGATGGAACGCGCGTCAAGCTGCAAATCTGGGAC 1572

Qy        297 ACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCC 356
              || ||||| ||||| || ||||| ||||| ||||| || |||||  | || || || ||
Db       1573 ACAGCTGGTCAGGAGCGATTCCGGAGCGTTACCCACGCCTATTATCGCGACGCGCACGCT 1632

Qy        357 TTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTC 416
              | || |||||||| ||| | ||||||||  |  | |  ||||||||  | ||||||||
Db       1633 CTACTGCTGCTGTACGACGTGACCAACAAGACCACCTATGACAACATTCGCGCCTGGCTG 1692

Qy        417 ACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCG 476
              |||||  |  ||||| || |||  ||||||||| ||| |  |  ||||||||||||
Db       1693 GGCGAGATCCGGGAGTACGCGCAGGAGGACGTGGTCATCGTTTTAATAGGCAACAAGGCC 1752

Qy        477 GA---TATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAG 533
              ||    |   |||||||  |   | |        || || || |||   ||||   |||||
Db       1753 GACTGCAGCGGCAGCGAGCGGCAGGTGAAGCGGGAGGATGGGGAGCGTTTGGGGCGGGAG 1812

Qy        534 TACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTT 593
              ||   || |||||| ||||||||   |||||||| ||  | |||||||||  |  ||||
Db       1813 CACAACGTGCCCTTCATGGAGACCTCGGCCAAGACGGGACTCAATGTGGAGCTGTCCTTC 1872

Qy        594 CTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTC 653
              ||  | |||| | |||| ||    || | |    | || |  |||||    |  |||
Db       1873 ACAGCGGTGGCCAGGCAACTAAAGAGTCGCGGCTACGAGCACGGCGATGATGGAAAGTTC 1932

Qy        654 CAGATCCGAGACTATGT 670
              |  | |  || | |||
Db       1933 AATGTGCATGATTTTGT 1949


RESULT 15
AAS27040
ID   AAS27040 standard; cDNA; 964 BP.
XX
AC   AAS27040;
XX
DT   07-NOV-2001  (first entry)
XX
DE   cDNA encoding novel signal transduction pathway protein, Seq ID 75.
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW   antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis C; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01312.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
```

PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR P-PSDB; AAU17123.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 75; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Query Match 7.8%; Score 209.4; DB 22; Length 964;
Best Local Similarity 69.3%; Pred. No. 5.4e-31;
Matches 285; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Qy 291 TGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCT 350
Db 3 TGGGACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCT 62

Qy 351 CAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCC 410
Db 63 CATGCTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCC 122

Qy 411 TGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAAC 470
Db 123 TGGCTGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAAC 182

Qy 471 AAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGG 530
Db 183 AAGGTGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAG 242

Qy 531 GAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCC 590

```
Db        243 GAGTATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCC 302

Qy        591 TTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGC 650
              ||    ||||| || |||||  ||||  | ||  ||  | |      |    |||||  ||
Db        303 TTCACAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGC 362

Qy        651 TTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              ||||  |  |  |  || || ||  ||    ||        |  ||  |||||||
Db        363 TTCCGGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 413
```

Search completed: June 23, 2003, 21:56:32
Job time : 570 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:32:44 ; Search time 3647 Seconds
(without alignments)
11874.618 Million cell updates/sec

Title: US-09-817-199B-1
Perfect score: 2674
Sequence: 1 ttcgcctgcgggccggcact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	921.6	34.5	1014	14	BM921365	BM921365 AGENCOURT
	2	884.2	33.1	887	9	AL522282	AL522282 AL522282
c	3	847.4	31.7	855	9	AL580781	AL580781 AL580781
c	4	802.6	30.0	833	9	AL522281	AL522281 AL522281
	5	755.4	28.2	881	12	BG759655	BG759655 602713348
	6	741.6	27.7	868	14	BQ690583	BQ690583 AGENCOURT
	7	733.2	27.4	761	13	BI819064	BI819064 603033417
c	8	702	26.3	702	14	BM981939	BM981939 UI-CF-EN1
	9	701.2	26.2	1051	12	BG283602	BG283602 602407930
c	10	696.8	26.1	700	14	BQ028124	BQ028124 UI-H-CO0-
	11	643.6	24.1	1069	12	BG282782	BG282782 602405867
c	12	626.4	23.4	629	14	BQ183276	BQ183276 UI-H-EU0-
	13	605.4	22.6	607	9	AF188522	AF188522 AF188522
	14	601.6	22.5	740	13	BI767046	BI767046 603054267
	15	600.6	22.5	691	12	BG253976	BG253976 602366910
	16	585.6	21.9	749	9	AL559085	AL559085 AL559085
	17	581.2	21.7	886	14	BQ686932	BQ686932 AGENCOURT
	18	553.2	20.7	676	10	BB598938	BB598938 BB598938
	19	539.2	20.2	603	10	AW955318	AW955318 EST367388
	20	529	19.8	529	14	BM706070	BM706070 UI-E-DW0-
	21	527.6	19.7	651	10	BB633978	BB633978 BB633978
	22	516.8	19.3	655	10	BB635649	BB635649 BB635649
	23	514.6	19.2	522	13	BM151643	BM151643 TCBAP1E10
c	24	503.4	18.8	507	13	BM151381	BM151381 TCBAP1D86
	25	473.6	17.7	480	10	AW270880	AW270880 xs05e04.x
	26	464.4	17.4	493	9	AA324998	AA324998 EST28027
	27	460.6	17.2	488	13	BM149118	BM149118 TCAAP2E63
	28	442.4	16.5	676	10	BB206788	BB206788 BB206788
	29	440.4	16.5	766	14	BQ188402	BQ188402 UI-E-EJ1-
	30	423.8	15.8	553	12	BE755280	BE755280 209068 MA
	31	422.4	15.8	733	10	BB619625	BB619625 BB619625
	32	415.2	15.5	437	14	W36293	W36293 HFBEST-70 H
	33	412.4	15.4	793	12	BG256040	BG256040 602367337
c	34	409.4	15.3	434	9	AA806096	AA806096 oc25b06.s
c	35	399.4	14.9	401	10	AW768344	AW768344 hk73a07.x
	36	397.4	14.9	657	10	BB196489	BB196489 BB196489
c	37	385	14.4	401	9	AA631787	AA631787 np68e10.s
c	38	381	14.2	403	9	AA878511	AA878511 oj19b02.s
	39	380.4	14.2	455	10	AW464449	AW464449 BP230015B
	40	377.8	14.1	385	9	AI243836	AI243836 qh84d08.x
	41	374.2	14.0	380	10	BE241772	BE241772 TCAAP2E02
	42	362.8	13.6	579	10	BB621927	BB621927 BB621927
	43	358	13.4	368	13	BM152052	BM152052 TCBAP1E12
	44	347	13.0	348	9	AA310682	AA310682 EST181496
	45	345.6	12.9	404	13	BM149006	BM149006 TCAAP2E57

ALIGNMENTS

RESULT 1
BM921365
LOCUS BM921365 1014 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6626159 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752779 5', mRNA sequence.
ACCESSION BM921365
VERSION BM921365.1 GI:19371744
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12787 row: i column: 04
High quality sequence stop: 735.
FEATURES Location/Qualifiers
source 1. .1014

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752779"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 237 a 288 c 295 g 194 t
ORIGIN

Query Match 34.5%; Score 921.6; DB 14; Length 1014;
Best Local Similarity 96.7%; Pred. No. 5.2e-93;
Matches 973; Conservative 0; Mismatches 29; Indels 4; Gaps 3;

```
Qy          2 TCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGC 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGC 67

Qy         62 CGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGA 127

Qy        122 CCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGAT 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGAT 187

Qy        182 CCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTT 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTT 247

Qy        242 CAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGC 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGC 307

Qy        302 TGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCT 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCT 367

Qy        362 TCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGA 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGA 427

Qy        422 GATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATAT 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATAT 487

Qy        482 GAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGT 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGT 547

Qy        542 TCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 TCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT 607

Qy        602 CGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCG 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 CGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCG 667

Qy        662 AGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGG 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGG 727

Qy        722 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCT 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCT 787
```

```
Qy        782 TATTCCAAGAGGCTGAGCCAAT-GGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 840
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        788 TATTCCAAGAGGCTGAGCCAATGGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 847

Qy        841 CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCAT--CCAC 898
              ||||||| |||||||||||||||||||||||| | ||||||    |||||||  | | |
Db        848 CTAGCAGGAAGCTATACTCCAACTCCTACTTGAGGTTCTGCGGGTCTCCCGCATTCCACC 907

Qy        899 AGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAA-GG 957
              || |||||||||||| |||||||||||||||| || ||| || |||||||||||| ||
Db        908 GGGAAGGGTAAAACACCTAGCTTTTATTTTAATAGGACCTAAATTTATACCAAAAAAGGG 967

Qy        958 CGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTT 1003
              |||||||||||||||||||| |||   || |    || ||  ||
Db        968 CGCCTGGATCCCCAAAAAACCCAGGGCTGGGGAGCTAGGGCCCCTT 1013
```

RESULT 2
AL522282
LOCUS AL522282 887 bp mRNA linear EST 13-FEB-2001
DEFINITION AL522282 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 5 prime, mRNA sequence.
ACCESSION AL522282
VERSION AL522282.1 GI:12785775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB008YF08"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 203 a 256 c 257 g 168 t 3 others
ORIGIN

Query Match 33.1%; Score 884.2; DB 9; Length 887;
Best Local Similarity 99.5%; Pred. No. 7.4e-89;
Matches 883; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy         15 GGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 74
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACTGCTCACCTCTCKTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 60

Qy         75 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 120

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194
```

```
Db        121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 180

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240

Qy        255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 GTGACTGTGGATGGCGTGAAAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 300

Qy        315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        301 TTCCGAAGCGTCACCCATGCTTATTACAGARATGCTCAGGCCTTGCTTCTGCTGTATGAC 360

Qy        375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 420

Qy        435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        421 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAARA 480

Qy        495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 540

Qy        555 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 600

Qy        615 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 660

Qy        675 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 720

Qy        735 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 780

Qy        795 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 840

Qy        855 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGG 901
              |||||||||||||||||||||||||||||||||||||||||||||
Db        841 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGG 887


RESULT 3
AL580781/c
LOCUS       AL580781                855 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL580781 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ014YB11 3
            prime, mRNA sequence.
ACCESSION   AL580781
VERSION     AL580781.1  GI:12947137
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 855)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .855
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ014YB11"
                     /clone_lib="LTI_NFL008_TC2"
                     /sex="male"
                     /tissue_type="T cells from T cell leukemia"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime  end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      204 a    224 c    224 g    201 t      2 others
ORIGIN

  Query Match           31.7%;  Score 847.4;  DB 9;  Length 855;
  Best Local Similarity   99.6%;  Pred. No. 8.5e-85;
  Matches  848;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1700 ATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTC 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 ATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTC 796

Qy       1760 AAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGC 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 AAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGC 736

Qy       1820 CCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTC 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTC 676

Qy       1880 TTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTT 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 TTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTT 616

Qy       1940 TCCTAGGTTTGGAATGTGTTGTGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGG 1999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 TCCTAGGTTTGGAATGTGTTGTGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGG 556

Qy       2000 GAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCC 2059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 GAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCC 496

Qy       2060 TCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGC 2119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGC 436

Qy       2120 TTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAA 2179
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        435 TTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGCTCACCCATTTCCGAGTTAA 376

Qy       2180 ACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTA 2239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 ACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTA 316

Qy       2240 GCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGA 2299
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        315 GCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCCGATCCCTGCCATGGGA 256

Qy       2300 AGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCT 2359
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        255 AGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGYTCTGCCGGCCT 196

Qy       2360 GGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGC 2419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGC 136

Qy       2420 TTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAA 2479
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAA 76

Qy       2480 GTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTG 2539
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTG 16

Qy       2540 GGGCCCAAGGC 2550
                ||||||||||||
Db         15 GGGCCCAAGGC 5


RESULT 4
AL522281/c
LOCUS       AL522281                833 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL522281 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 3
            prime, mRNA sequence.
ACCESSION   AL522281
VERSION     AL522281.1  GI:12785774
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB008YF08"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      206 a    207 c    215 g    200 t      5 others
ORIGIN

  Query Match             30.0%;  Score 802.6;  DB 9;  Length 833;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-80;
  Matches  814;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       1673 GTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGC 1732
                :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        825 KTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGC 766

Qy       1733 CCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTAC 1792
                |||||||||||||||||||||||||||||||||||||:||||||||:||||||||||
Db        765 CCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGSCCTCGGCCYCCCAAAGTAC 706

Qy       1793 TGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGG 1852
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        705 TGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGG 646

Qy       1853 CAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGT 1912
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 CAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGT 586

Qy       1913 TAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGA 1972
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGA 526

Qy       1973 GAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGT 2032
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 GAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGT 466

Qy       2033 ATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTC 2092
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 ATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTC 406

Qy       2093 TCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTG-AGTCAGGTTAGGT 2151
                |||||||||||||||||||||||||||||||||||| |  |||| | ||||||||||||
Db        405 TCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCACAAACTGCACTCAGGTTAGGT 346

Qy       2152 GGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGT 2211
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 GGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGT 286

Qy       2212 GGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATC 2271
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 GGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATC 226

Qy       2272 TCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCA 2331
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 TCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCA 166

Qy       2332 GGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGC 2391
                ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        165 GGCATCTTTACTGCAGMTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGC 106

Qy       2392 TGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 2451
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 TGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 46

Qy       2452 GCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCC 2496
                |||||||||||||||||||||||||||||||||||||||||| |||||:
Db         45 GCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGCCCTCCV 1


RESULT 5
BG759655
LOCUS       BG759655                881 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602713348F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853626 5',
            mRNA sequence.
ACCESSION   BG759655
VERSION     BG759655.1  GI:14070308
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 881)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1699  row: p  column: 11
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:4853626"
                     /clone_lib="NIH_MGC_48"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      220 a    236 c    261 g    164 t
ORIGIN

  Query Match             28.2%;  Score 755.4;  DB 12;  Length 881;
  Best Local Similarity   94.6%;  Pred. No. 1.1e-74;
  Matches  805;  Conservative    0;  Mismatches   41;  Indels    5;  Gaps    2;

Qy        566 GACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGC 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGC 61

Qy        626 CGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAA 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAA 121

Qy        686 GCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGC 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGC 181

Qy        746 ACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGG 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGG 241

Qy        806 GGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTC 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTC 301

Qy        866 CTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTAT 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTAT 361

Qy        926 TTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTG 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTG 421

Qy        986 GGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATA 1045
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        422 GGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCAGCCCTCCCTCCTAAGCATA 481

Qy       1046 ACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGG 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGG 541

Qy       1106 GCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGC 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGC 601

Qy       1166 CTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCC 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCC 661

Qy       1226 CTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGG 1285
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        662 CTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGAGGGGAGGGG 721

Qy       1286 GAGGACAAGGGGCAGAGAGTAGGGTCTA---GCTGGCTATCTCTGGCCTTACTAACACCC 1342
              ||||| ||||||||||||||| ||    |      |    |||||||| ||| || | |
Db        722 GAGGAAAAGGGGCAGAGAGTAAGGGACAAGCATGGCAAGTCTCTGGCCGTACAAAGAGCG 781

Qy       1343 CCCTGG--AGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATAT 1400
```

```
              |||||   |||||||| | |   |||| |||      |  | ||||||||||||||||
Db        782 CCCTGGGAGGGCATGCCCTTATAAGCCAGGACAACGAGCCCATTGGGCACCTGGAAATAG 841

Qy       1401 TGGTTCCAGGC 1411
              |||  |  | |
Db        842 TGGACCGGGAC 852


RESULT 6
BQ690583
LOCUS       BQ690583                868 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8186311 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251642
            5', mRNA sequence.
ACCESSION   BQ690583
VERSION     BQ690583.1  GI:21815899
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2396  row: k  column: 03
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6251642"
                     /clone_lib="NIH_MGC_110"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of  California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      215 a    239 c    240 g    172 t      2 others
ORIGIN

  Query Match             27.7%;  Score 741.6;  DB 14;  Length 868;
  Best Local Similarity   93.6%;  Pred. No. 3.5e-73;
  Matches  805;  Conservative    0;  Mismatches   51;  Indels    4;  Gaps    3;

Qy        490 AAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCC 60

Qy        550 TGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGG 120

Qy        610 AACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATG 180

Qy        670 TAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAG 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAG 240
```

```
Qy        730 AGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAA 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAA 300

Qy        790 GAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGG 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGG 360

Qy        850 AGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAA 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAA 420

Qy        910 AACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCC 480

Qy        970 CAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTT-GGGGGGCCGGC 1028
              |||||||||||||||||||||||||||||||||||||  | |||| ||||||||||||
Db        481 CAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCAGGNACTTGGGGGGGCCGGC 540

Qy       1029 CCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGAT 1088
              ||||||||||||||||  ||  ||||||||||||||||||||||||||||||||||||||
Db        541 CCTCCCTCCTAAGCATACNAAGAGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGAT 600

Qy       1089 GCACTTT-GGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGA 1147
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCACTTTGGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGA 660

Qy       1148 GGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCAT 1207
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCAT 720

Qy       1208 CTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAG 1267
              |||||  |||||||||||||||||||||||||||||||  |   ||| ||   ||||
Db        721 CTCCCCAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAGCCTCAAGGAAAATAAAGAA 780

Qy       1268 ATATGGAGATGGGAGGGGGAGGACAAGGGG--CAGAGAGTAGGGTCTAGCTGGCTATCTC 1325
                    |  ||| ||||||| |||||||||   | | || |||||||| |   |   | ||
Db        781 TATGAAAAGTGGAAGGGGGAAGACAAGGGGGCAAAAAAGAAGGGTCTAACCTGGCTTATC 840

Qy       1326 TGGCCTTACTAACACCCCCC 1345
              |      ||||  | |||
Db        841 TCTTGGCCTTAACTACACCC 860
```

RESULT 7
BI819064

```
LOCUS       BI819064                761 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603033417F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174308 5',
            mRNA sequence.
ACCESSION   BI819064
VERSION     BI819064.1  GI:15930614
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 761)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11434  row: b  column: 05
            High quality sequence stop: 759.
```

```
FEATURES             Location/Qualifiers
     source          1. .761
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5174308"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      174 a    219 c    204 g    164 t
ORIGIN
```

```
  Query Match           27.4%;  Score 733.2;  DB 13;  Length 761;
  Best Local Similarity   99.3%;  Pred. No. 3.3e-72;
  Matches  757;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy        728 AGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCC 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCC 60

Qy        788 AAGAGGCTGAGCCAATG-GGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCA 846
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAGGCTGAGCCAATGTGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCA 120

Qy        847 GGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGG 180

Qy        907 TAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGAT 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGAT 240

Qy        967 CCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCA 300

Qy       1027 GCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAG 360

Qy       1087 ATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAG 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAG 420

Qy       1147 AGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCA 480

Qy       1207 TCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGA 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGA 540

Qy       1267 GATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCT 600

Qy       1327 GGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGG 1386
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCTTACTAACAACCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGG 660

Qy       1387 GCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCC 1446
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| ||
Db        661 GCACCTGGAAATATGGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAG-CC 719

Qy       1447 CTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 1488
```

```
Db        720 CTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 761

RESULT 8
BM981939/c
LOCUS       BM981939                 702 bp    mRNA    linear   EST 21-MAR-2002
DEFINITION  UI-CF-EN1-adg-h-11-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-adg-h-11-0-UI 3', mRNA sequence.
ACCESSION   BM981939
VERSION     BM981939.1  GI:19604935
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-adg-h-11-0-UI"
                     /clone_lib="UI-CF-EN1"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_LIB=UI-CF-EN1
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_SEQ=CTGCTCAGGT"
BASE COUNT      166 a    181 c    171 g    184 t
ORIGIN

  Query Match           26.3%;  Score 702;  DB 14;  Length 702;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-69;
  Matches  702;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1953 ATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 ATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAA 643

Qy       2013 GCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCC 2072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCC 583

Qy       2073 ATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTC 2132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 ATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTC 523

Qy       2133 CTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATAT 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATAT 463

Qy       2193 GAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTG 2252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 GAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTG 403

Qy       2253 TCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAA 2312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 TCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAA 343

Qy       2313 CATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGG 2372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGG 283

Qy       2373 GGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAA 2432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAA 223

Qy       2433 ATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCC 2492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 ATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCC 163

Qy       2493 TCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAG 2552
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 TCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAG 103

Qy       2553 GTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAA 2612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAA 43

Qy       2613 TTAAACTGCAGTATTTTGGAAAGCAAAAAAAAAAAAAAAAAA 2654
             ||||||||||||||||||||||||||||||||||||||||||
Db          42 TTAAACTGCAGTATTTTGGAAAGCAAAAAAAAAAAAAAAAA 1

RESULT 9
BG283602
LOCUS       BG283602                 1051 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602407930F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520191 5',
            mRNA sequence.
ACCESSION   BG283602
VERSION     BG283602.1  GI:13033710
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
          Plate: LLAM10417  row: k  column: 08
          High quality sequence stop: 782.
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4520191"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      275 a    287 c    314 g    175 t
ORIGIN

  Query Match             26.2%;  Score 701.2;  DB 12;  Length 1051;
  Best Local Similarity   99.3%;  Pred. No. 8.3e-69;
  Matches  725;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy         40 ACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCC 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCC 60

Qy        100 CGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCG 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCG 120

Qy        160 TCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCA 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCA 180

Qy        220 TAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGA 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGA 240

Qy        280 AGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATT 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATT 300

Qy        340 ACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACA 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACA 360

Qy        400 ACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGC 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGC 420

Qy        460 TGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGA 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGA 480

Qy        520 CCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATG 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATG 540

Qy        580 TGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGG 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGG 600

Qy        640 ATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAG-AAGAAGCGCTCCAGCTGC 698
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        601 ATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAAGAAGCGCTCCAGCTGC 660

Qy        699 TGCTCCTTCATGTGAATCCCAGGGGG-CAGAGAGGAGGCTCTGGAGGCACACAGGATGCA 757
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        661 TGCTCCTTCATGTGAATCCCAGGGGGCCAGAGAGGAGGCTCTGGAGGCACACAGGATGCA 720

Qy        758 GCCTTCCCCC 767
              || | ||| |
Db        721 GCTTCCCCTC 730
```

```
RESULT 10
BQ028124/c
LOCUS       BQ028124                 700 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-H-CO0-ari-d-07-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
            IMAGE:3106595 3', mRNA sequence.
ACCESSION   BQ028124
VERSION     BQ028124.1  GI:19763403
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3106595"
                     /clone_lib="NCI_CGAP_Sub9"
                     /tissue_type="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
                     Cholonic mucosa with Crohns disease, Cholonic mucosa with
                     ulcerative colitis, Fetal thymus, Cervix, Cervical
                     adenosquamous carcinoma, Ligament cells, Prostate
                     carcinoma, Bladder carcinoma, Brain oligodenroga ;
                     NCI_CGAP_Sub9 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are CGTC, AACG, GGGCC,
                     GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
                     information, contact: Bento Soares, bento-soares@uiowa.edu
                     TAG_LIB=UI-H-CO0
                     TAG_TISSUE=Fetal thymus
                     TAG_SEQ=AACG"
BASE COUNT      166 a    182 c    170 g    182 t
ORIGIN

  Query Match             26.1%;  Score 696.8;  DB 14;  Length 700;
  Best Local Similarity   99.7%;  Pred. No. 3.6e-68;
  Matches  698;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1950 GGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGAT 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGAT 641

Qy       2010 TAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCA 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCA 581
```

```
Qy       2070 CCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCC 2129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCC 521

Qy       2130 CTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 2189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 CTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 461

Qy       2190 TATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCAT 2249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 TATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCAT 401

Qy       2250 GTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGG 2309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGG 341

Qy       2310 AAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAG 2369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 AAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAG 281

Qy       2370 AGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCA 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 AGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCA 221

Qy       2430 GAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTG 2489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTG 161

Qy       2490 GCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGG 2549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGG 101

Qy       2550 CAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCA 2609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 CAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCA 41

Qy       2610 GAATTAAACTGCAGTATTTTGGAAAGCAAAAAAAAAAAAA 2649
             ||||||||||||||||||||||||| ||||||||||||||
Db         40 GAATTAAACTGCAGTATTTTGGAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 11
BG282782
LOCUS       BG282782                1069 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602405867F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4518136 5',
            mRNA sequence.
ACCESSION   BG282782
VERSION     BG282782.1  GI:13031999
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1069)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10412  row: e  column: 17
            High quality sequence stop: 689.
FEATURES             Location/Qualifiers
     source          1. .1069
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4518136"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      264 a    296 c    295 g    214 t
ORIGIN
```

```
  Query Match            24.1%;  Score 643.6;  DB 12;  Length 1069;
  Best Local Similarity  94.3%;  Pred. No. 1.7e-62;
  Matches  722;  Conservative    0;  Mismatches   39;  Indels    5;  Gaps    5;

Qy       1104 GGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCA 1163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCA 60

Qy       1164 GCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGG 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGG 120

Qy       1224 CCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGG 1283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGG 180

Qy       1284 GGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCC 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCC 240

Qy       1344 CCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGG 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGG 300

Qy       1404 TTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCC 1463
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        301 TTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCT-CCCCCCCTGAATCC 359

Qy       1464 CTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGA 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 CTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGA 419

Qy       1524 GTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTA 1583
             |||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GTTAGGTTG-TATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTA 478

Qy       1584 GTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCA 1643
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||
Db        479 GTGCAATCTCCGCTCACTACAACCTCCAATCCCTGGGGCTCAAGCGATCCTCCCAACTCA 538

Qy       1644 GCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTT 1703
             ||| |||||||||||||||||||||||||||||||||||||||||||||||| ||| |
Db        539 GCCCCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTAGTA-TA 597

Qy       1704 TTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGC 1763
                |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||
Db        598 AAGGTAGACACAGGGTTTCGCCATGTAGCCCAGGCTGGTCTTGAGTTCCTGAGCTCAAGC 657

Qy       1764 AACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAG-AAGGCACCATGCCCA 1822
             || ||||||| |||| |||||||||||||||||| || ||| | ||||||| |||
Db        658 AA-CTGCCGGACTCGGGCTCCCAAAGTACTGGGATGACCCGCCGAAAGGCACATAGCCAG 716

Qy       1823 GGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCA 1868
             |||||||       ||  || |  | |||||||| |||||| || |
Db        717 GGCTAGAATGTGACTAATCCCAATCGTTGGCAGGGATGCAGGTCAA 762
```

```
RESULT 12
BQ183276/c
LOCUS       BQ183276                629 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-H-EU0-azr-e-05-0-UI.s1 NCI_CGAP_Car1 Homo sapiens cDNA clone
```

```
            IMAGE: 5852524 3', mRNA sequence.
ACCESSION   BQ183276
VERSION     BQ183276.1  GI:20358826
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5852524"
                     /clone_lib="NCI_CGAP_Car1"
                     /tissue_type="Osteoarthritic Cartilage"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Organ: Knee; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Car1 is a cDNA library containing the following
                     tissue(s): Osteoarthritic Cartilage The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     TGATCACGCT.
                     TAG_LIB=UI-H-EU0
                     TAG_TISSUE=osteoarthritic cartilage
                     TAG_SEQ=TGATCACGCT"
BASE COUNT      150 a    162 c    156 g    160 t      1 others
ORIGIN

  Query Match             23.4%;  Score 626.4;  DB 14;  Length 629;
  Best Local Similarity   99.7%;  Pred. No. 2.1e-60;
  Matches  627;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2026 GACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTT 2085
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        629 GACATGTATCCCTTTGACCCCAAGCTNTGCCTCCTCCCTGACCACCCATGCCCTTTCCTT 570

Qy       2086 TAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGG 2145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 TAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGG 510

Qy       2146 TTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAG 2205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAG 450

Qy       2206 TCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCA 2265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCA 390

Qy       2266 CATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGA 2325
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        389 CATATCTCCCAAAGCTCCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGA 330

Qy       2326 CCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGT 2385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGT 270

Qy       2386 ATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAG 2445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAG 210

Qy       2446 CAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGA 2505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGA 150

Qy       2506 GCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATC 2565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATC 90

Qy       2566 CAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTA 2625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          89 CAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTA 30

Qy       2626 TTTTGGAAAGCAAAAAAAAAAAAAAAAAA 2654
              ||||||||||||||||||||||||||||||
Db          29 TTTTGGAAAGCAAAAAAAAAAAAAAAAAA 1


RESULT 13
AF188522
LOCUS       AF188522               607 bp    mRNA    linear   EST 02-MAR-2000
DEFINITION  AF188522 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens cDNA clone
            ISG 3, mRNA sequence.
ACCESSION   AF188522
VERSION     AF188522.1  GI:7144568
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 607)
  AUTHORS   Ye,Z. and Connor,J.R.
  TITLE     Identification of Iron Regulated Genes by Rescreening cDNA
            Libraries from SSH with Antisense Probe from Three Iron Conditions
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Ye Z
            Neuroscience and Anatomy
            Pennsylvania State University College of Medicine
            500 University Drive, Hershey, PA 17033, USA
            library screened by SSH and reverse Northern blot; increased
            expression in iron loading was confirmed by Northern blot.
FEATURES             Location/Qualifiers
     source          1. .607
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ISG 3"
                     /clone_lib="Homo sapiens ATCC HTB-12; SW1088"
                     /tissue_type="astrocytoma"
                     /cell_line="ATCC HTB-12; SW1088"
                     /note="Organ: liver"
BASE COUNT      154 a    153 c    155 g    145 t
ORIGIN

  Query Match             22.6%;  Score 605.4;  DB 9;  Length 607;
  Best Local Similarity   99.8%;  Pred. No. 4.4e-58;
  Matches  606;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1794 GGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGC 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGC 60

Qy       1854 AGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTT 1913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 AGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTT 120

Qy       1914 AAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAG 1973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAG 180

Qy       1974 AAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTA 2033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTA 240

Qy       2034 TCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCT 2093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCT 300

Qy       2094 CAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGG 2153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGG 360

Qy       2154 TGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGG 2213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGG 420

Qy       2214 GTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTC 2273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTC 480

Qy       2274 CCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGG 2333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGG 540

Qy       2334 CATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTG 2393
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        541 CATCTTTACTGCAGCTCTGCCGGCCTGGAGGAGGAGAGGGGGAGGAAGAAGTATGCGCTG 600

Qy       2394 CACATTT 2400
              |||||||
Db        601 CACATTT 607
```

RESULT 14
BI767046

```
LOCUS       BI767046                740 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603054267F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203600 5',
            mRNA sequence.
ACCESSION   BI767046
VERSION     BI767046.1  GI:15758624
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11510  row: f  column: 17
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5203600"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      171 a    198 c    222 g    149 t
ORIGIN
```

```
  Query Match             22.5%;  Score 601.6;  DB 13;  Length 740;
  Best Local Similarity   96.2%;  Pred. No. 9.8e-58;
  Matches  712;  Conservative    0;  Mismatches   19;  Indels    9;  Gaps    9;

Qy         42 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 60

Qy        102 CCCTGCAGTCC-GAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGT 160
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTGCAGTCCTGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGT 120

Qy        161 CGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCAT 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCAT 180

Qy        221 AGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAA 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAA 240

Qy        281 GCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTA 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTA 300

Qy        341 CAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAA 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAA 360

Qy        401 CATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGG-TGATCATGC 459
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db        361 CATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTTGATCATGC 420

Qy        460 TGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGA 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGA 480

Qy        520 CCTTGGCCAGGGAGTACGGTGTTCCCTTCCT-GGAGACCAGCGCCAAGAC-TGGCATGAA 577
              |||||| |||| |||||||||||||||||| ||||| ||| |||||||| |||||||||
Db        481 CCTTGGGCAGGTAGTACGGTGTTCCCTTCCTGGGAGAACAGTGCCAAGACTTGGCATGAA 540

Qy        578 TGTGGAGTTA-GCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 636
              |||||||||| || ||||| | |||||||| ||||||  |||||||||||||||||||||
Db        541 TGTGGAGTTAGGCTTTTCTTGGCATCGCCAGGGAACTTGAATACCGGGCCGGGCATCAGG 600

Qy        637 CGGATGAGCCCAGCTTCCAGA-TCCGAGACTATGTAG-AGTCCCAGAAGAAGCGCTCCAG 694
              ||||||| |||||||||||||| |||||||||||||| |||||||||||||||||||| ||
Db        601 CGGATGAACCCAGCTTCCAGATTCCGAGACTATGTAGAAGTCCCAGAAGAAGCGCTCAAG 660

Qy        695 C-TGCTGCTCCTTCATGTGAA-TCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGG 752
              | ||||| |||||||||||||| |||||||||||||||||||||||||||| || ||||||
Db        661 CTTGCTGGTCCTTCATGTGAAGTCCCAGGGGGCAGAGAGGAGGCTCTGGACGCCCACAGG 720

Qy        753 ATGCAGCCTTCCCCCTCCCA 772
              |||||||||| | ||  |||
Db        721 ATGCAGCCTTTCTCCATCCA 740
```

RESULT 15
BG253976

```
LOCUS       BG253976                691 bp    mRNA    linear   EST 13-FEB-2001
```

```
DEFINITION  602366910F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475067 5',
            mRNA sequence.
ACCESSION   BG253976
VERSION     BG253976.1  GI:12763792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10300  row: c  column: 04
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4475067"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      175 a    185 c    190 g    141 t
ORIGIN

  Query Match             22.5%;  Score 600.6;  DB 12;  Length 691;
  Best Local Similarity   97.6%;  Pred. No. 1.3e-57;
  Matches  663;  Conservative    0;  Mismatches    9;  Indels    7;  Gaps    5;

Qy        287 GATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGA 346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGA 73

Qy        347 TGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAG 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAG 133

Qy        407 GGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGG 466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGG 193

Qy        467 CAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGC 253

Qy        527 CAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 CAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTT 313

Qy        587 AGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCC 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 AGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCC 373

Qy        647 CAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTT 433

Qy        707 CATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        434 CATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAG-CTTCCCC 492

Qy        767 CTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGA--GAAAGATGGAGGACTC 824
              ||||||||||||||||||||||||||||||||||||||||||   || ||||||||||||
Db        493 CTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGACGAAACGATGGAGGACTC 552

Qy        825 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGC-GG 883
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        553 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCAGG 612

Qy        884 TCTCCCCGCATCCACAGGGA-GGGTAAAACACTTAGCTTTTATTTT--AATAGTACATAA 940
              |||||||||||||||||||| |||||||||||||||||||||||| |  |   ||||||||
Db        613 TCTCCCCGCATCCACAGGGAGGGGTAAAACACTTAGCTTTTATTCTCAATACGTACATAA 672

Qy        941 TTTAATACCAAAAAAGGCG 959
              |||||||||||||| | | |
Db        673 TTTAATACCAAAACAAGGG 691


Search completed: June 24, 2003, 00:50:48
Job time : 3651 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:35:29 ; Search time 121 Seconds
(without alignments)
6777.303 Million cell updates/sec

Title: US-09-817-199B-1
Perfect score: 2674
Sequence: 1 ttcgcctgcgggccggcact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2607.8	97.5	2612	4	US-09-484-970B-142	Sequence 142, App
	2	873	32.6	875	4	US-09-075-454-10	Sequence 10, Appl
	3	292.6	10.9	1340	2	US-08-824-873-2	Sequence 2, Appli
	4	292.6	10.9	1340	3	US-09-198-184-2	Sequence 2, Appli
c	5	193.4	7.2	99500	4	US-09-798-096-10	Sequence 10, Appl
c	6	193.2	7.2	87350	3	US-08-781-891-79	Sequence 79, Appl
c	7	193.2	7.2	87543	4	US-09-791-211-3	Sequence 3, Appli
c	8	192.8	7.2	8453	4	US-09-167-681-45	Sequence 45, Appl
c	9	192.4	7.2	112132	4	US-09-741-150-3	Sequence 3, Appli
	10	192.2	7.2	6678	3	US-08-816-617A-1	Sequence 1, Appli
	11	192.2	7.2	9301	4	US-09-449-218D-18	Sequence 18, Appl
c	12	192.2	7.2	35060	3	US-08-814-095-7	Sequence 7, Appli
	13	192.2	7.2	98844	4	US-09-791-211-10	Sequence 10, Appl
c	14	191.4	7.2	29629	4	US-09-729-995-3	Sequence 3, Appli
	15	190.8	7.1	6769	1	US-08-480-784-20	Sequence 20, Appl
	16	190.8	7.1	6769	1	US-08-483-553-20	Sequence 20, Appl
	17	190.8	7.1	6769	1	US-08-487-002-20	Sequence 20, Appl
	18	190.8	7.1	6769	1	US-08-483-554B-20	Sequence 20, Appl
	19	190.8	7.1	6769	1	US-08-488-011B-20	Sequence 20, Appl
	20	190.8	7.1	6769	4	US-08-850-727-20	Sequence 20, Appl
	21	190.8	7.1	6769	5	PCT-US95-10202-20	Sequence 20, Appl
	22	190.8	7.1	6769	5	PCT-US95-10203-20	Sequence 20, Appl
	23	190.8	7.1	6769	5	PCT-US95-10220-20	Sequence 20, Appl
c	24	190.4	7.1	45546	4	US-09-146-053-6	Sequence 6, Appli
	25	190	7.1	45716	4	US-08-965-048-5	Sequence 5, Appli
	26	190	7.1	45989	4	US-08-965-048-6	Sequence 6, Appli
	27	189	7.1	14581	4	US-08-520-373D-4	Sequence 4, Appli
	28	189	7.1	22481	4	US-08-367-841A-43	Sequence 43, Appl
	29	189	7.1	22481	5	PCT-US95-07201-43	Sequence 43, Appl
	30	189	7.1	22484	4	US-09-875-223-2	Sequence 2, Appli
c	31	189	7.1	55827	4	US-09-813-133A-3	Sequence 3, Appli
	32	189	7.1	99500	4	US-09-798-096-10	Sequence 10, Appl
c	33	188.4	7.0	36651	4	US-09-738-894A-3	Sequence 3, Appli
	34	188.4	7.0	59065	4	US-09-813-817-3	Sequence 3, Appli
	35	188.4	7.0	59065	4	US-09-978-197-3	Sequence 3, Appli
c	36	187.8	7.0	50000	4	US-09-146-053-3	Sequence 3, Appli
	37	187.6	7.0	8133	4	US-09-659-791A-10	Sequence 10, Appl
	38	187	7.0	3867	4	US-09-347-114A-81	Sequence 81, Appl
c	39	187	7.0	43950	4	US-09-735-934A-3	Sequence 3, Appli
c	40	186.6	7.0	8353	3	US-08-611-587-1	Sequence 1, Appli
c	41	186.4	7.0	1043	4	US-09-165-868-4	Sequence 4, Appli
	42	186.4	7.0	3694	4	US-09-232-200-46	Sequence 46, Appl
	43	186.4	7.0	3694	4	US-09-232-197-46	Sequence 46, Appl
	44	186.4	7.0	3694	4	US-09-232-201-46	Sequence 46, Appl
	45	186.4	7.0	3704	4	US-09-232-200-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-484-970B-142
; Sequence 142, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 142
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1
US-09-484-970B-142

Query Match 97.5%; Score 2607.8; DB 4; Length 2612;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2609; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          27 CTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCC 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCC 60

Qy          87 CCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTG 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTG 120

Qy         147 GGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTG 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTG 180

Qy         207 TCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGAT 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGAT 240

Qy         267 GGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTC 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTC 300

Qy         327 ACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAA 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAA 360
```

```
Qy    387 TCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGAC 446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGAC 420

Qy    447 GTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCC 506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCC 480

Qy    507 GAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAG 566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAG 540

Qy    567 ACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCC 626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCC 600

Qy    627 GGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAG 686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAG 660

Qy    687 CGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCA 746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCA 720

Qy    747 CACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGG 806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGG 780

Qy    807 GAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCC 866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCC 840

Qy    867 TACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATT 926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATT 900

Qy    927 TTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGG 986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGG 960

Qy    987 GAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAA 1046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAA 1020

Qy   1047 CAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGG 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGG 1080

Qy   1107 CAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCC 1166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAGGTAATGACTCCATCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCC 1140

Qy   1167 TTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCC 1226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCC 1200

Qy   1227 TTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGG 1286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGG 1260

Qy   1287 AGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCT 1346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCT 1320

Qy   1347 GGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTC 1406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTC 1380

Qy   1407 CAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCCTGAATCCCTG 1466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCCTGAATCCCTG 1440

Qy   1467 GCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTT 1526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTT 1500

Qy   1527 TTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTG 1586
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTG 1560

Qy   1587 CAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCC 1646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCC 1620

Qy   1647 GCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTT 1706
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTT 1680

Qy   1707 GTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAAC 1766
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAAC 1740

Qy   1767 CTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCT 1826
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCT 1800

Qy   1827 AGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGC 1886
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGATGTGTCTTATCCCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGC 1860

Qy   1887 AGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGG 1946
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGG 1920

Qy   1947 TTTGGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAA 2006
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TTTGGAATGTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAA 1980

Qy   2007 GATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGA 2066
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GATTAAGCAAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGA 2040

Qy   2067 CCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCT 2126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCT 2100

Qy   2127 CCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATG 2186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATG 2160

Qy   2187 CAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTT 2246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTT 2220

Qy   2247 CATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGAC 2306
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2221 CATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCCGATCCCTGCCATGGGAAGTGGAC 2280

Qy   2307 AGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGG 2366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGG 2340

Qy   2367 GAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAG 2426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAG 2400

Qy   2427 GCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCC 2486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCC 2460

Qy   2487 GTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCA 2546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCA 2520

Qy   2547 AGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTT 2606
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTT 2580

Qy       2607 TCAGAATTAAACTGCAGTATTTTGGAAAGCA 2637
                 |||||||||||||||||||||||||||||||
Db       2581 TCAGAATTAAACTGCAGTATTTTGGAAAGCA 2611


RESULT 2
US-09-075-454-10
; Sequence 10, Application US/09075454
; Patent No. 6391580
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Patterson, Chandra
;    APPLICANT:  Batra, Sajeev
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  RAS PROTEINS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Word Perfect 6.1/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/075,454
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/766,551
;      FILING DATE:  DECEMBER 12, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0168-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  875 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  UCMCL5T01
;      CLONE:  1528559
US-09-075-454-10

  Query Match             32.6%;  Score 873;  DB 4;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-186;
  Matches  873;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 60

Qy        106 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 120

Qy        166 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 180

Qy        226 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 240

Qy        286 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 300

Qy        346 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 360

Qy        406 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 420

Qy        466 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 480

Qy        526 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 540

Qy        586 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 600

Qy        646 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 660

Qy        706 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 720

Qy        766 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 780

Qy        826 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 840

Qy        886 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 918
              ||||||||||||||||||||||||||||||||
Db        841 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 873


RESULT 3
US-08-824-873-2
; Sequence 2, Application US/08824873
; Patent No. 5843717
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
```

```
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/824,873
;       FILING DATE:  Filed Herewith
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J.
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0240 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-855-0555
;       TELEFAX:  415-845-4166
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1340 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  PANCNOT04
;       CLONE:  738957
US-08-824-873-2

  Query Match           10.9%;  Score 292.6;  DB 2;  Length 1340;
  Best Local Similarity   70.9%;  Pred. No. 1.5e-56;
  Matches  416;  Conservative    0;  Mismatches  169;  Indels    2;  Gaps    2;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              ||||||| || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db         21 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 80

Qy        176 CCTGA-TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCA 234
              |||  | | |||||| || || || |||||| | || ||||||||  ||||||| |||
Db         81 GCTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139

Qy        235 TAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGG 294
              | |||||| ||||||| ||  ||   |||||||| |||| ||||||||||||| ||||
Db        140 TTGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGG 199

Qy        295 ACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGG 354
              |||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| |
Db        200 ACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATG 259

Qy        355 CCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGC 414
              |  |||| ||||| || ||  |||||||||  | || || |||||||||  ||||||||
Db        260 CTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGC 319

Qy        415 TCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGG 474
              | || ||||| || ||||| ||||||   ||||||| | |||||||||| || |||||||
Db        320 TGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGG 379

Qy        475 CGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGT 534
              |||      |   ||  | ||| |       || ||||| |||   |||||| |||||
Db        380 TGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGT 439

Qy        535 ACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTC 594
              | ||  | |||||| |||||||||||||||||| ||| | || ||||| || |||||
Db        440 ATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCA 499

Qy        595 TGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCC 654
              |||||| || ||||| |||| | ||  ||  | |      |    ||||| ||||||
Db        500 CAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCC 559

Qy        655 AGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              | | |  || || ||  ||    ||      |  ||  |||||||
Db        560 GGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606


RESULT 4
US-09-198-184-2
; Sequence 2, Application US/09198184
; Patent No. 6010859
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Guegler, Karl
;    TITLE OF INVENTION:  NOVEL RAB PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/198,184
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/824,873
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0240 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1340 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PANCNOT04
;      CLONE:  738957
US-09-198-184-2

  Query Match           10.9%;  Score 292.6;  DB 3;  Length 1340;
  Best Local Similarity   70.9%;  Pred. No. 1.5e-56;
  Matches  416;  Conservative    0;  Mismatches  169;  Indels    2;  Gaps    2;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              ||||||| || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db         21 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 80

Qy        176 CCTGA-TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCA 234
              |||  | | |||||| || || || |||||| | || ||||||||  ||||||| |||
Db         81 GCTGGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139

Qy        235 TAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGG 294
              | |||||| ||||||| ||  ||   |||||||| |||| ||||||||||||| ||||
Db        140 TTGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGG 199

Qy        295 ACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGG 354
              |||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| |
Db        200 ACACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATG 259

Qy        355 CCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGC 414
              |  |||| ||||| || ||  |||||||||  | || || |||||||||  ||||||||
Db        260 CTCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGC 319

Qy        415 TCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGG 474
              | || ||||| || ||||| ||||||   ||||||| | |||||||||| || |||||||
Db        320 TGACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGG 379

Qy        475 CGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGT 534
```

```
Db        380 TGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAAGCTGGCCAAGGAGT 439

Qy        535 ACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTC 594

Db        440 ATGGACTGCCCTTCATGGAGACCAGCGCCAAGACGGGCCTCAACGTGGACTTGGCCTTCA 499

Qy        595 TGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCC 654

Db        500 CAGCCATAGCAAAGGAGTTGAAGCAGCGCTCCATGAAGGCTCCCAGCGAGCCGCGCTTCC 559

Qy        655 AGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701

Db        560 GGCTGCATGATTACGTTAAGAGGGAGGGTCGAGGGGCCTCCTGCTGC 606


RESULT 5
US-09-798-096-10/c
; Sequence 10, Application US/09798096
; Patent No. 6399378
; GENERAL INFORMATION:
;  APPLICANT: Donna T. Ward
;  APPLICANT:  Andrew T. Watt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF RECQL2 EXPRESSION
;  FILE REFERENCE: RTS-0207
;  CURRENT APPLICATION NUMBER: US/09/798,096
;  CURRENT FILING DATE:  2001-03-01
;  NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
;   LENGTH: 99500
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-798-096-10

  Query Match             7.2%;  Score 193.4;  DB 4;  Length 99500;
  Best Local Similarity   79.5%;  Pred. No. 9.4e-34;
  Matches  241;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584

Db      66170 TTTGTTTTGTTTTGTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGG 66111

Qy       1585 TGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAG 1644

Db      66110 CACAATCTTGGCTCACTACAACCTCTGCCTGCT-GGGTTCAAGCCATTCTCCTGCCTCAG 66052

Qy       1645 CCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTT 1704

Db      66051 CCTCCCAAGTAGCTGGGATTATGGATGCCCACCATCGCGCCCAGCTAATTTTTGTATTTT 65992

Qy       1705 TTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764

Db      65991 TAGTAGAGACAGAGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGCA 65932

Qy       1765 ACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGG 1824

Db      65931 ATCCACTTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCGGGCCCAGC 65872

Qy       1825 CTA 1827

Db      65871 CTA 65869


RESULT 6
US-08-781-891-79/c
; Sequence 79, Application US/08781891
; Patent No. 6090620
;  GENERAL INFORMATION:
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Yu, Chang-En
;    APPLICANT:  Oshima, Junko
;    APPLICANT:  Mulligan, John T.
;    APPLICANT:  Schellenberg, Gerald D.
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
;    TITLE OF INVENTION:  WERNER'S SYNDROME
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,891
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6090620tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  240052.419
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  87350 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-79

  Query Match             7.2%;  Score 193.2;  DB 3;  Length 87350;
  Best Local Similarity   79.9%;  Pred. No. 1e-33;
  Matches  239;  Conservative    0;  Mismatches   59;  Indels    1;  Gaps    1;

Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584

Db      42283 TTTTTTTTTTTTTTTTTTNGAGACAGAGTCTCGCTCTGTCTCCCAGGCTGAAGTGCAGTGG 42224

Qy       1585 TGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAG 1644

Db      42223 CACAATCTTGACTCACTGCAACCTCCACCTCCT-GGGTTCAAGCAATTCTCCTGCCTCAG 42165

Qy       1645 CCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTT 1704

Db      42164 CCTCCTGAATAGCTGGGATTACAGGCACCCAACACCACGCCCAGCTAATTTTTGTACTTT 42105

Qy       1705 TTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764

Db      42104 TAGTAGAAACGGAGTTTCATCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTG 42045

Qy       1765 ACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAG 1823

Db      42044 ATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCAG 41986


RESULT 7
US-09-791-211-3/c
; Sequence 3, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
;  APPLICANT:  Donna T. Ward
;  APPLICANT:  Andrew T. Watt
;  TITLE OF INVENTION:  ANTISENSE MODULATION OF WRN EXPRESSION
;  FILE REFERENCE:  RTS-0205
;  CURRENT APPLICATION NUMBER: US/09/791,211
;  CURRENT FILING DATE:  2001-02-23
;  NUMBER OF SEQ ID NOS:  90
; SEQ ID NO 3
```

; LENGTH: 87543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 7421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7427
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 11609
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12605
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12742
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29370
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29979
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29980
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29981
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30136
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30140
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31205
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31206
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31592
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33095
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33160
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34066
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34072
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 36816
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 39020
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42164
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42459
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46808
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46823
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46826
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 47291
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52786
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52787
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 53384
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 54684
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59215
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59235
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59242
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 63290
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 66614
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68660
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68697
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68718
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68733
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68739
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 69785
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 79134
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 79198
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 86336
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
US-09-791-211-3

Query Match 7.2%; Score 193.2; DB 4; Length 87543;
Best Local Similarity 79.9%; Pred. No. 1e-33;
Matches 239; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

```
Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584
              |||| ||||| |||||| || |  |||||||| |||||| |||||||||| |||||||
Db      42476 TTTTTTTTTTTTTTTTTNGAGACAGAGTCTCGCTCTGTCTCCCAGGCTGAAGTGCAGTGG 42417

Qy       1585 TGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAG 1644
              ||||||   |||||| |||||||||| ||| ||| |||||| || ||||  ||||||
Db      42416 CACAATCTTGACTCACTGCAACCTCCACCTCCT-GGGTTCAAGCAATTCTCCTGCCTCAG 42358

Qy       1645 CCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTT 1704
              || ||  | |||||||||| || |||     | || ||| ||  |||||||||||| |||
Db      42357 CCTCCTGAATAGCTGGGATTACAGGCACCCAACACCACGCCCAGCTAATTTTTGTACTTT 42298

Qy       1705 TTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764
              | ||||| || | |||||  ||||||| |||||||||||||||| |||||| ||||||
Db      42297 TAGTAGAAACGGAGTTTCATCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTG 42238

Qy       1765 ACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAG 1823
              | ||||| |||||||||||||||||| |||||||||| ||   || ||||||||||||
Db      42237 ATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCAG 42179


RESULT 8
US-09-167-681-45/c
; Sequence 45, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
;  APPLICANT: Weinshilboum, M.D., Richard M.
;  APPLICANT: Raftogianis, Rebecca B.
;  APPLICANT: Wood, Thomas C.
;  APPLICANT: Otterness, Diane M.
;  TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
;  FILE REFERENCE: 07039/118001
;  CURRENT APPLICATION NUMBER: US/09/167,681A
;  CURRENT FILING DATE: 1998-10-07
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 8447
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (4361)...(4507)
;   NAME/KEY: CDS
;   LOCATION: (4612)...(4737)
;   NAME/KEY: CDS
;   LOCATION: (4827)...(4925)
;   NAME/KEY: CDS
;   LOCATION: (6322)...(6447)
;   NAME/KEY: CDS
;   LOCATION: (6543)...(6638)
;   NAME/KEY: CDS
;   LOCATION: (7137)...(7316)
;   NAME/KEY: CDS
;   LOCATION: (7439)...(7553)
US-09-167-681-45

  Query Match             7.2%;  Score 192.8;  DB 4;  Length 8453;
  Best Local Similarity   75.3%;  Pred. No. 5.8e-34;
  Matches  253;  Conservative    0;  Mismatches   82;  Indels    1;  Gaps    1;

Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584
              ||||  |||| ||| | ||| |||||| |||| ||| ||||||||||||   |||||||
Db       2562 TTTTTCTTTTTTTTTATGTGAGATGGAGACTCGCTCTCTCGCCCAGGCTGCAATGCAGTAG 2503

Qy       1585 TGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAG 1644
                | ||||||||||||  ||||||| |  ||| ||| |||||||||| || |  ||||||
Db       2502 CACGATCTCCGCTCACTGTAACCTCCCCCTCCT-GGGTTCAAGCGATTCTTCTGCCTCAG 2444

Qy       1645 CCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTT 1704
              || ||  |||||||||| || |||||   |||| |||| | || ||||||||||||||
Db       2443 CCTCCCGAGTAGCTGGGATTACAGGTGCACACCACCACATCCGGTTAATTTTTGTATTTT 2384

Qy       1705 TTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764
              | ||||| || |||||  |||||||| |||||||||||   || ||||||||||
Db       2383 TAGTAGAGACGGGGTTTTGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGAGCTCAGATG 2324

Qy       1765 ACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGG 1824
              | | ||| |||||||||||||||||| ||| ||||||| |||  || |||||||||  |
Db       2323 ATCCGCCCGCCTCGGCCTCCCAAAGTGCTGAGATTACAGGCATGAGCCACCATGCCTGGC 2264

Qy       1825 CTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATG 1860
              | | | |  | | |       ||   |  ||| | |
Db       2263 CGACAAGCTTTCTCTAATTGACCCAGGCAAGGGAGG 2228


RESULT 9
US-09-741-150-3/c
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL000968
;  CURRENT APPLICATION NUMBER: US/09/741,150
;  CURRENT FILING DATE:  2000-12-21
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 112132
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(112132)
;   OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

  Query Match             7.2%;  Score 192.4;  DB 4;  Length 112132;
  Best Local Similarity   78.4%;  Pred. No. 1.6e-33;
  Matches  243;  Conservative    0;  Mismatches   66;  Indels    1;  Gaps    1;

Qy       1526 TTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGT 1585
              |||  |||||||||||||| |  | |||| |||||||| |||||||||   |||||| |
Db      97233 TTTTATTTTATTTTTTTGAGACAGGGTCTTGTTCTGTCACCCAGGCTGGAATGCAGTGGC 97174

Qy       1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGC 1645
              | ||||  ||||| | ||||||||  || |   |||||||| ||||||||||||||||||
Db      97173 GTGATCTTGGCTCATTGCAACCTCC-GTCTCCCAGGCTCAAGTGATCCTCCCACCTCAGC 97115

Qy       1646 CGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTT 1705
              | || |||||| || ||||| |||  |   ||| ||||||  ||||||||||||||||||
Db      97114 CTCCCAAGTAGTTGTGACTACAGGCATAAGCCACCACACCCAGCTAATTTTTGTATTTTT 97055

Qy       1706 TGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAA 1765
              |||||| |  || ||| ||| ||| ||||||||||||||||||| |||    |||||||||
Db      97054 TGTAGAGAAGGGATTTTGCCGTGTCGCCCAGGCTGGTCTTGAACTCCCAGACTCAAGCAA 96995

Qy       1766 CCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGC 1825
                |  || ||||| |||||||||||| |||||||||| |||  || ||||  |||||| |
Db      96994 TCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCGCGCCCAGCC 96935

Qy       1826 TAGATGTGTC 1835
              | ||| | ||
Db      96934 TGGATATTTC 96925


RESULT 10
US-08-816-617A-1
; Sequence 1, Application US/08816617A
; Patent No. 6022741
;  GENERAL INFORMATION:
;    APPLICANT:  Ting, Jenny P.-Y.
;    APPLICANT:  Piskurich, Janet
```

```
;    TITLE OF INVENTION:  No. 6022741el Regulatory Genetic DNA that
;    TITLE OF INVENTION:  Regulates the Class II Transactivator (CIITA)
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bell, Seltzer, Park & Gibson
;      STREET:  1211 East Morehead Street
;      CITY:  Charlotte
;      STATE:  No. 6022741th Carolina
;      COUNTRY:  United States
;      ZIP:  28234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/816,617A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sibley, Kenneth D.
;      REGISTRATION NUMBER:  31,665
;      REFERENCE/DOCKET NUMBER:  5470-143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-420-2200
;      TELEFAX:  919-881-3175
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6678 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-816-617A-1

  Query Match             7.2%;  Score 192.2;  DB 3;  Length 6678;
  Best Local Similarity   77.8%;  Pred. No. 7.3e-34;
  Matches  245;  Conservative    0;  Mismatches   68;  Indels    2;  Gaps    1;

Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584
              |||||||||  ||||||||||| |||||||| ||||  |||||||||. |||||||  |
Db       3043 TTTTGTTTTTGATTTTTTGAAATAGAGTCTCGCTCTGAAGCCCAGGCTGGAGTGCAGTGG 3102

Qy       1585 TGCAATCTCCGCTCACT--ACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTC 1642
              ||  ||||| |||||||    | |  |  |||||   || |||||| || ||||  ||||
Db       3103 TGTGATCTCGGCTCACTTGCAACCTCCACCTCCCCCAGGTTCAAGCAATTCTCCTGCCTC 3162

Qy       1643 AGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATT 1702
              |||| || ||||| |||||| || ||| ||| |||| |||||||||| ||||||||  ||||
Db       3163 AGCCTCCCAAGTAACTGGGATTACAGGCGTGCACCACCACACCTGGATAATTTTTCTATT 3222

Qy       1703 TTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAG 1762
              ||| ||||| |||||||||| |||||||| |||||||||. | ||||||||||  |||| |
Db       3223 TTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTTTCGAATTCCTGACCTCAGG 3282

Qy       1763 CAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCA 1822
                | | ||| ||||| | |||| |||||  || ||||||| ||   || |||||  ||||
Db       3283 TGATCCGCCCGCCTCAGTCTCCTAAAGTGTTGAGATTACAGGCGTGAGCCACCACGCCCG 3342

Qy       1823 GGCTAGATGTGTCTT 1837
              | ||  | | |||||
Db       3343 GCCTGCAGGAGTCTT 3357


RESULT 11
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary E.
;  APPLICANT:  Galas, David J.
;  APPLICANT:  Kovacevich, Brian
;  APPLICANT:  Mulligan, John T.
;  APPLICANT:  Paeper, Bryan W.
;  APPLICANT:  Van Ness, Jeffrey
;  APPLICANT:  Winkler, David G.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
;  TITLE OF INVENTION:  BONE MINERALIZATION
;  FILE REFERENCE: 240083.508
;  CURRENT APPLICATION NUMBER: US/09/449,218D
;  CURRENT FILING DATE:  1999-11-24
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 9301
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-449-218D-18

  Query Match             7.2%;  Score 192.2;  DB 4;  Length 9301;
  Best Local Similarity   77.4%;  Pred. No. 8.1e-34;
  Matches  233;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy       1525 TTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAG 1584
              | ||  |||  || ||||||  |  |||||||| ||| | ||||||||||  |||||||
Db       6477 TGTTTGTTTGTTTGTTTTGAGACAGAGTCTCGCTCTATTGCCCAGGCTGGAGTGCAGTGT 6536

Qy       1585 TGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAG 1644
                |||||||  ||| ||| |||| ||  |   |  ||  ||||| ||| ||||  ||||||
Db       6537 CACAATCTCGGCTTACTGCAACTTCTGCCTTCCCGGATTCAAGTGATTCTCCTGCCTCAG 6596

Qy       1645 CCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTT 1704
              || ||  |||||||||||  || ||||| || ||| ||| ||||||||||||||||||||
Db       6597 CCTCCAGAGTAGCTGGGATTACAGGTGCGTGCCACCACGCCTGGCTAATTTTTGTATTTT 6656

Qy       1705 TTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCA 1764
              |  |||| || |||||||| |||||||| ||||||| |||| ||| || ||| ||||||
Db       6657 TGATAGAGACGGGGTTTCACCATGTTGGCCAGGCTAGTCTCGAACTCTTGACCTCAAGTG 6716

Qy       1765 ACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGG 1824
              | ||||| |||||||||||||||||| |||||||||| ||   || |||||  |||||
Db       6717 ATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACACCCAGC 6776

Qy       1825 C 1825
              |
Db       6777 C 6777


RESULT 12
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
;  GENERAL INFORMATION:
;    APPLICANT:  Soreq, Hermona
;    APPLICANT:  Zakut, Haim
;    APPLICANT:  Shani, Moshe
;    TITLE OF INVENTION:  TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;    TITLE OF INVENTION:  ANTI-CHOLINESTERASE SUBSTNACES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  KOHN & ASSOCIATES
;      STREET:  30500 No. 6025183thwestern Highway, Suite 410
;      CITY:  Farmington Hills
;      STATE:  Michigan
;      COUNTRY:  U.S.
;      ZIP:  48334
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,095
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Montgomery, Ilene N.
;       REGISTRATION NUMBER:  38,972
;       REFERENCE/DOCKET NUMBER:  2391.00066
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (248) 539-5050
;       TELEFAX:  (248) 539-5055
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  35060 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  other nucleic acid
;       DESCRIPTION:  /desc = "Cosmid including ACHE
;       DESCRIPTION:  promotor, ACHE gene and ARS gene"
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:  7q22
;     FEATURE:
;       NAME/KEY:  promoter
;       LOCATION:  4089..22464
;       OTHER INFORMATION:  /function= "ACHE Promotor"
;       OTHER INFORMATION:  /standard_name= "ACHE Promotor"
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  22465..22537
;       OTHER INFORMATION:  /function= "non-translated"
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 1
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  24090..25177
;       IDENTIFICATION METHOD:  experimental
;       OTHER INFORMATION:  /function= "(translation start:
;       OTHER INFORMATION:  24110)"
;       OTHER INFORMATION:  /evidence= EXPERIMENTAL
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 2
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  25524..26009
;       IDENTIFICATION METHOD:  experimental
;       OTHER INFORMATION:  /evidence= EXPERIMENTAL
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 3
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  27005..27274
;       IDENTIFICATION METHOD:  experimental
;       OTHER INFORMATION:  /evidence= EXPERIMENTAL
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 4
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  27255..28007
;       IDENTIFICATION METHOD:  experimental
;       OTHER INFORMATION:  /evidence= EXPERIMENTAL
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 5
;     FEATURE:
;       NAME/KEY:  terminator
;       LOCATION:  27385..27387
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  28008..28129
;       IDENTIFICATION METHOD:  experimental
;       OTHER INFORMATION:  /evidence= EXPERIMENTAL
;       OTHER INFORMATION:  /gene= "ACHE"
;       OTHER INFORMATION:  /number= 6
;     FEATURE:
;       NAME/KEY:  terminator
;       LOCATION:  28129..28131
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (34528..34895)
;       OTHER INFORMATION:  /function= "arsenite resistance
;       OTHER INFORMATION:  gene"
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 1
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (34092..34358)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 2
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33779..33963)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 3
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33493..33591)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 4
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33297..33408)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 5
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32959..33094)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 6
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32569..32628)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 7
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32386..32468)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 8
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31894..32080)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 9
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31363..31534)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 10
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31131..31284)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 11
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30816..31011)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 12
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30470..30626)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 13
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30187..30274)
;       OTHER INFORMATION:  /gene= "AR"
```

```
;      OTHER INFORMATION:  /number= 14
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  complement (29945..30073)
;      OTHER INFORMATION:  /gene= "AR"
;      OTHER INFORMATION:  /number= 15
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  complement (29664..29856)
;      OTHER INFORMATION:  /gene= "ARS"
;      OTHER INFORMATION:  /number= 16
US-08-814-095-7

  Query Match             7.2%;  Score 192.2;  DB 3;  Length 35060;
  Best Local Similarity   77.4%;  Pred. No. 1.3e-33;
  Matches  233;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy       1530 TTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAA 1589
             ||||| ||||||| | | ||||| | | |||||| |||||||||  ||||||| | |  |
Db      14160 TTTTTTTTTTTTTTAGACGGAGTTTAGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGTGA 14101

Qy       1590 TCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCC 1649
             |||||  |||||  ||||||| | |||  | | ||||||||| ||||  |||||||| ||
Db      14100 TCTCCAATCACTGTAACCTCCGCCCCCCAGAGTTCAAGCGATTCTCCTGCCTCAGCCTCC 14041

Qy       1650 GAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTA 1709
              ||||||||||||| || ||| |  | ||| ||||||||||||||||||  |||||| |||
Db      14040 CAAGTAGCTGGGATTACAGGCGCCTGCCACCACACCTGGCTAATTTTTCCATTTTTAGTA 13981

Qy       1710 GACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTG 1769
             ||  |||||||||  ||| ||| |||||||||||||||| |||||| ||||||  | |
Db      13980 GAGGCAGGGTTTCTGCATATTGGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCCA 13921

Qy       1770 CCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGA 1829
             || |||| ||| |||||||| || |||||||| ||| ||| |||||  ||  | |   |
Db      13920 CCTGCCTTGGCTTCCCAAAGTGCTAGGATTACAGGCATAAGCCACCACACCTGGCCCTCA 13861

Qy       1830 T 1830
              |
Db      13860 T 13860


RESULT 13
US-09-791-211-10
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
;  APPLICANT:  Donna T. Ward
;  APPLICANT:  Andrew T. Watt
;  TITLE OF INVENTION:  ANTISENSE MODULATION OF WRN EXPRESSION
;  FILE REFERENCE:  RTS-0205
;  CURRENT APPLICATION NUMBER: US/09/791,211
;  CURRENT FILING DATE:  2001-02-23
;  NUMBER OF SEQ ID NOS:  90
; SEQ ID NO 10
;   LENGTH: 98844
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 24962
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 64383
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65468
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65469
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65470
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65471
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 87130
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 89049
;   OTHER INFORMATION: unknown
;   OTHER INFORMATION:
US-09-791-211-10

  Query Match             7.2%;  Score 192.2;  DB 4;  Length 98844;
  Best Local Similarity   79.0%;  Pred. No. 1.7e-33;
  Matches  241;  Conservative    0;  Mismatches   63;  Indels    1;  Gaps    1;

Qy       1526 TTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGT 1585
              ||| ||||| |||||||| |||| |||||  || || |||||||| |  ||||||| ||
Db       2480 TTTTTTTTTTTTTTTTTTGAGATGGCGTCTCACTCCGTTGCCCAGGCAGGAGTGCAGTGGT 2539

Qy       1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGC 1645
              || ||||| ||||||| |||||||| |  |||  || |||||||||| ||||  ||||||||
Db       2540 GCCATCTCAGCTCACTGCAACCTCCGCCTCCT-AGGTTCAAGCGATTCTCCTGCCTCAGC 2598

Qy       1646 CGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTT 1705
              |  |  | |||||||||  || ||| | || ||| ||  |||||||||||||| ||||||||
Db       2599 CTTCTGACTAGCTGGGATTACAGGAGCGTGCCACCATGCCTGGCTAATTTTGGTATTTTT 2658

Qy       1706 TGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAA 1765
              ||||| || | ||||  |||||||  |||||||||||||||| |||||| || |||  |
Db       2659 AGTAGAGACGGAGTTTAACCATGTTGGTCAGGCTGGTCTTGAACTCCTGAACTTAAGTGA 2718

Qy       1766 CCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGC 1825
              | ||| |||| |||||||||||| |||||||||| ||| || ||||  ||||| |
Db       2719 TCCGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCACACCCAGCC 2778

Qy       1826 TAGAT 1830
              || ||
Db       2779 TAAAT 2783


RESULT 14
US-09-729-995-3/c
; Sequence 3, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000904
;  CURRENT APPLICATION NUMBER: US/09/729,995
;  CURRENT FILING DATE:  2000-12-06
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 29629
;   TYPE: DNA
;   ORGANISM: Human
US-09-729-995-3

  Query Match             7.2%;  Score 191.4;  DB 4;  Length 29629;
  Best Local Similarity   76.3%;  Pred. No. 1.8e-33;
  Matches  248;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1524 GTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTA 1583
              ||||| ||||| |||||||| |  |||||||  |||||| |||||||||  ||||| |
Db      11930 GTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCACTCTGTCCCCCAGGCTGGAGTGCAATG 11871

Qy       1584 GTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCA 1643
              | |  ||||| ||||||| || |||||||| | |   || ||||| ||| ||||  |||||
Db      11870 GCGTGATCTCGGCTCACTGCAGCCTCCAC-CTCCCAGGTTCAAGAGATTCTCCTGCCTCA 11812
```

```
Qy       1644 GCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTT 1703
              ||| ||  ||||| ||||| || |||||| | ||| ||| || ||||| ||||||||||.
Db      11811 GCCTCCCGAGTAGTTGGGATTACAGGTGTCTGCCACCACGCCCGGCTACTTTTTGTATTT 11752

Qy       1704 TTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGC 1763
              || ||||| |||||||||||||||||||  |||||||||||||| |||||| |||| |
Db      11751 TTAGTAGAGACAGGGTTTCGCCATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGT 11692

Qy       1764 AACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAG 1823
                | |  ||  |||||||||||||||| |||||||||| ||   || ||||   |||||
Db      11691 GATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCCCCAG 11632

Qy       1824 GCTAGATGTGTCTTATCCCAATCCT 1848
                ||   |  || | || |  ||
Db      11631 CTGGGAGCAGAGTTCTTCCTAGGCT 11607


RESULT 15
US-08-480-784-20
; Sequence 20, Application US/08480784
; Patent No. 5693473
;  GENERAL INFORMATION:
;    APPLICANT:  Skolnick, Mark H.
;    APPLICANT:  Goldgar, David E.
;    APPLICANT:  Miki, Yoshio
;    APPLICANT:  Swenson, Jeff
;    APPLICANT:  Kamb, Alexander
;    APPLICANT:  Harshman, Keith D.
;    APPLICANT:  Shattuck-Eidens, Donna M.
;    APPLICANT:  Tavtigian, Sean V.
;    APPLICANT:  Wiseman, Roger W.
;    APPLICANT:  Futreal, P. Andrew
;    TITLE OF INVENTION:  17q-Linked Breast and Ovarian Cancer
;    TITLE OF INVENTION:  Susceptibility Gene
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Venable, Baetjer, Howard & Civiletti, LLP
;      STREET:  1201 New York Avenue, N.W., Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,784
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/409,305
;      FILING DATE:  24-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/348,824
;      FILING DATE:  29-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/308,104
;      FILING DATE:  16-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/300,266
;      FILING DATE:  02-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/289,221
;      FILING DATE:  12-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  24884-109347
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-962-4810
;      TELEFAX:  202-962-8300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6769 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-480-784-20

  Query Match             7.1%;  Score 190.8;  DB 1;  Length 6769;
  Best Local Similarity   75.5%;  Pred. No. 1.5e-33;
  Matches  237;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1526 TTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGT 1585
              ||||||| | ||||||||| |||| |||||  ||||| ||||||||||  |||||||  |
Db       6242 TTTGTTTGTTTTTTTTTGAGATGGGGTCTCACTCTGTTGCCCAGGCTGGAGTGCAGTGAT 6301

Qy       1586 GCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGC 1645
              |||||||  |||||||  |||||||||||  ||  |||||||||  |||||||||| |||||
Db       6302 GCAATCTTGGCTCACTGCAACCTCCACCTCCAAAGGCTCAAGCTATCCTCCCACTTCAGC 6361

Qy       1646 CGCCGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTT 1705
              | ||  |||||||||||||| |||  |  | |||  ||||||  ||  |||||||      |||
Db       6362 CTCCTGAGTAGCTGGGACTACAGGCGCATGCCACCACACCCGGTTAATTTTTTGTGGTTT 6421

Qy       1706 TGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAA 1765
              | ||||  |  |||||||  |||||||  ||  ||||||||||  ||  |||||  ||||||||
Db       6422 TATAGAGATGGGGTTTCACCATGTTACCGAGGCTGGTCTCAAACTCCTGGACTCAAGCAG 6481

Qy       1766 CCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGC 1825
              |||||  | || ||||||||||||  |||   |||||  ||   ||  |||  ||||  | |
Db       6482 TCTGCCCACTTCAGCCTCCCAAAGTGCTGCAGTTACAGGCTTGAGCCACTGTGCCTGGCC 6541

Qy       1826 TAGATGTGTCTTAT 1839
              |      |  |||  |
Db       6542 TGCCCTTTACTTTT 6555


Search completed: June 24, 2003, 00:53:00
Job time : 124 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 21:56:39 ; Search time 376 Seconds
(without alignments)
10435.905 Million cell updates/sec

Title: US-09-817-199B-1
Perfect score: 2674
Sequence: 1 ttcgcctgcgggccggcact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2674	100.0	2674	10	US-09-817-199A-1	Sequence 1, Appli
	2	2581.6	96.5	2623	9	US-09-764-868-71	Sequence 71, Appl
	3	2031.8	76.0	13182	10	US-09-817-199A-3	Sequence 3, Appli
	4	1286	48.1	1316	9	US-09-764-868-493	Sequence 493, App
	5	1099.6	41.1	1106	9	US-10-037-270-959	Sequence 959, App
	6	1086	40.6	1116	10	US-09-794-257-13	Sequence 13, Appl
	7	873	32.6	875	12	US-10-051-986-10	Sequence 10, Appl
	8	683.4	25.6	24707	9	US-10-274-968-3	Sequence 3, Appli
	9	683.4	25.6	24707	10	US-09-740-027-3	Sequence 3, Appli
	10	600.6	22.5	601	10	US-09-817-199A-24	Sequence 24, Appl
	11	576	21.5	576	10	US-09-794-257-15	Sequence 15, Appl
	12	573	21.4	601	10	US-09-817-199A-25	Sequence 25, Appl
	13	443.4	16.6	447	10	US-09-867-550-1811	Sequence 1811, Ap
	14	415	15.5	417	9	US-09-918-995-34999	Sequence 34999, A
	15	233.4	8.7	434	9	US-09-918-995-648	Sequence 648, App
	16	209.4	7.8	964	9	US-09-764-868-75	Sequence 75, Appl
	17	204.8	7.7	10859	9	US-09-764-891-7408	Sequence 7408, Ap
	18	203.6	7.6	36221	9	US-09-954-556-29	Sequence 29, Appl
c	19	200.8	7.5	1691139	9	US-10-067-514-1	Sequence 1, Appli
c	20	200.6	7.5	21833	10	US-09-764-877-2275	Sequence 2275, Ap
c	21	199.8	7.5	1130	10	US-09-764-877-3373	Sequence 3373, Ap
	22	198.6	7.4	15745	10	US-09-764-877-3236	Sequence 3236, Ap
c	23	198.6	7.4	17216	10	US-09-764-877-3565	Sequence 3565, Ap
c	24	198.6	7.4	17217	10	US-09-764-877-3566	Sequence 3566, Ap
c	25	198.4	7.4	4962	9	US-10-091-504-1902	Sequence 1902, Ap
c	26	198.4	7.4	4962	9	US-10-091-504-1903	Sequence 1903, Ap
c	27	198.4	7.4	4962	10	US-09-764-869-1902	Sequence 1902, Ap
c	28	198.4	7.4	4962	10	US-09-764-869-1903	Sequence 1903, Ap
c	29	198	7.4	32205	9	US-09-764-891-10213	Sequence 10213, A
c	30	198	7.4	32205	9	US-10-205-428-1012	Sequence 1012, Ap
c	31	197.6	7.4	7791	10	US-09-764-877-3700	Sequence 3700, Ap
c	32	197.4	7.4	65464	9	US-09-859-888-3	Sequence 3, Appli
	33	197.2	7.4	2754	10	US-09-745-763-133	Sequence 133, App
	34	197.2	7.4	15515	10	US-09-822-860-3	Sequence 3, Appli
c	35	196.8	7.4	17335	9	US-10-092-154-1280	Sequence 1280, Ap
c	36	196.8	7.4	17335	10	US-09-764-847-1280	Sequence 1280, Ap
c	37	196.8	7.4	19882	9	US-10-092-154-1281	Sequence 1281, Ap
c	38	196.8	7.4	19882	10	US-09-764-847-1281	Sequence 1281, Ap
	39	196.6	7.4	13821	10	US-09-764-877-2595	Sequence 2595, Ap
c	40	196.2	7.3	7960	9	US-10-091-504-2327	Sequence 2327, Ap
c	41	196.2	7.3	7960	10	US-09-764-869-2327	Sequence 2327, Ap
	42	196.2	7.3	19167	9	US-09-764-891-8028	Sequence 8028, Ap
	43	195.8	7.3	11960	9	US-09-764-891-8501	Sequence 8501, Ap
	44	195.8	7.3	203654	10	US-09-820-905-3	Sequence 3, Appli
c	45	195.4	7.3	30175	10	US-09-738-878-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-817-199A-1
; Sequence 1, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001187
; CURRENT APPLICATION NUMBER: US/09/817,199A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
US-09-817-199A-1

Query Match 100.0%; Score 2674; DB 10; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TTCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCGCCTGCGGGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCG 60

Qy         61 CCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACG 120

Qy        121 ACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGA 180

Qy        181 TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACT 240

Qy        241 TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG 300
Qy        301 CTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGC 360
Db        301 CTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGC 360
Qy        361 TTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTG 420
Db        361 TTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTG 420
Qy        421 AGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATA 480
Db        421 AGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATA 480
Qy        481 TGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTG 540
Db        481 TGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTG 540
Qy        541 TTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA 600
Db        541 TTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA 600
Qy        601 TCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCC 660
Db        601 TCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCC 660
Qy        661 GAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAG 720
Db        661 GAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAG 720
Qy        721 GGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGC 780
Db        721 GGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGC 780
Qy        781 TTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 840
Db        781 TTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTC 840
Qy        841 CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAG 900
Db        841 CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAG 900
Qy        901 GGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGC 960
Db        901 GGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGC 960
Qy        961 CTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGG 1020
Db        961 CTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGG 1020
Qy       1021 GGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGG 1080
Db       1021 GGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGG 1080
Qy       1081 ACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCT 1140
Db       1081 ACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACCCTCAGTTCAGCT 1140
Qy       1141 GGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTC 1200
Db       1141 GGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAGGAGCTTATCTTC 1200
Qy       1201 GCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAG 1260
Db       1201 GCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAG 1260
Qy       1261 ATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCT 1320
Db       1261 ATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCT 1320
Qy       1321 ATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCAC 1380
Db       1321 ATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCAC 1380
```

```
Qy       1381 ATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGG 1440
Db       1381 ATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGG 1440
Qy       1441 GAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAAC 1500
Db       1441 GAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAAC 1500
Qy       1501 CTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCT 1560
Db       1501 CTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCT 1560
Qy       1561 GTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGG 1620
Db       1561 GTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGG 1620
Qy       1621 GCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATC 1680
Db       1621 GCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTGTGTACCATC 1680
Qy       1681 ACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTG 1740
Db       1681 ACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTTGCCCAGGCTG 1740
Qy       1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTA 1800
Db       1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAAGTACTGGGATTA 1800
Qy       1801 CACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATG 1860
Db       1801 CACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCTTTGGCAGGCATG 1860
Qy       1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCC 1920
Db       1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGAGCC 1920
Qy       1921 AGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAAGAGAAATCCC 1980
Db       1921 AGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAAAAGAGAAATCCC 1980
Qy       1981 TGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTT 2040
Db       1981 TGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTT 2040
Qy       2041 GACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAG 2100
Db       2041 GACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAG 2100
Qy       2101 ATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGG 2160
Db       2101 ATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGG 2160
Qy       2161 TCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTC 2220
Db       2161 TCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTC 2220
Qy       2221 TGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGC 2280
Db       2221 TGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGC 2280
Qy       2281 TCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTT 2340
Db       2281 TCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTT 2340
Qy       2341 ACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTT 2400
Db       2341 ACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTT 2400
Qy       2401 CTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCA 2460
Db       2401 CTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCA 2460
```

```
Qy       2461 GTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTG 2520

Qy       2521 ACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAG 2580

Qy       2581 CTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGGAAAGCAAAA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTTTGGAAAGCAAAA 2640

Qy       2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
             ||||||||||||||||||||||||||||||||||
Db       2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674


RESULT 2
US-09-764-868-71
; Sequence 71, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PT232
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
;   LENGTH: 2623
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-868-71

  Query Match             96.5%;  Score 2581.6;  DB 9;  Length 2623;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2606;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy         42 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCG 71

Qy        102 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTC 131

Qy        162 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATA 191

Qy        222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 281
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        192 GCCACCGTCGGCATAGACTTCAGCAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 251

Qy        282 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTAC 311

Qy        342 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 371

Qy        402 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 ATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTG 431

Qy        462 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACC 491

Qy        522 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTG 551

Qy        582 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGAT 611

Qy        642 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGC 671

Qy        702 TCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCT 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 TCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCT 731

Qy        762 TCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGA 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 TCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGA 791

Qy        822 CTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGC 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 CTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGC 851

Qy        882 GGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAAT 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 GGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAAT 911

Qy        942 TTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTT 1001
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 TTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTT 971

Qy       1002 TTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGC 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 TTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGC 1031

Qy       1062 TCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCA 1121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 TCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCA 1091

Qy       1122 TCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGC 1181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 TCGCACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGC 1151

Qy       1182 TCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAA 1241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 TCTCCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAA 1211

Qy       1242 GACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGA 1301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 GACCAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGA 1271

Qy       1302 GAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACA-CCCCCCTGGAGGCATGCCCCT 1360
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       1272 GAGTAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCCTGGAGGCATGCCCCT 1331

Qy       1361 TTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCT 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 TTTCTCCAGCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCT 1391

Qy       1421 CTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTC 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 CTGGACTTCAGATCCTGGGGGAGCCCCTCCCCCCCCTGAATCCCTGGCTTAGCTACCTTC 1451

Qy       1481 CTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTT 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 CTGCCTGTGCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTT 1511

Qy       1541 TTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCAC 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 TTGAAATGGAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCAC 1571

Qy       1601 TACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGG 1660
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
```

```
Db       1572 TACAACCTCCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCCCCGAAGTAGCTGG 1631
Qy       1661 GACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTT 1720
Db       1632 GACTATAGGTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTT 1691
Qy       1721 TCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGC 1780
Db       1692 TCGCCATGTTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGC 1751
Qy       1781 CTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATC 1840
Db       1752 CTCCCAAAGTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATC 1811
Qy       1841 CCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTA 1900
Db       1812 CCAATCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTA 1871
Qy       1901 GCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTG 1960
Db       1872 GCCCTCCTGGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTG 1931
Qy       1961 TG-AAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 2019
Db       1932 TGAAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACC 1991
Qy       2020 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 2079
Db       1992 TCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCT 2051
Qy       2080 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 2139
Db       2052 TTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGA 2111
Qy       2140 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 2199
Db       2112 GTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAA 2171
Qy       2200 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 2259
Db       2172 ACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCT 2231
Qy       2260 TGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 2319
Db       2232 TGGTCACATATCTCCCAAAGCTCCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGG 2291
Qy       2320 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGA 2379
Db       2292 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGAGGAGAGGGGGAGGA 2351
Qy       2380 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 2439
Db       2352 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGC 2411
Qy       2440 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 2499
Db       2412 TCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCA 2471
Qy       2500 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 2559
Db       2472 GGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAG 2531
Qy       2560 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 2619
Db       2532 GAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACT 2591
Qy       2620 GCAGTATTTTGGAAAGCAAAAAAAAAAAAAA 2651
Db       2592 GCAGTATTTTGGAAAGCAAAAAAAAAAAAA 2623

RESULT 3
US-09-817-199A-3

; Sequence 3, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001187
;  CURRENT APPLICATION NUMBER: US/09/817,199A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 13182
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(13182)
;   OTHER INFORMATION: n = A,T,C or G
US-09-817-199A-3

  Query Match           76.0%;  Score 2031.8;  DB 10;  Length 13182;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 2039;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        589 CCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCA 648
Db      10060 CCATTGTCTCTTCTTCAAGGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGCCCA 10119
Qy        649 GCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCA 708
Db      10120 GCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCA 10179
Qy        709 TGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCT 768
Db      10180 TGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCCCT 10239
Qy        769 CCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTG 828
Db      10240 CCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCACTG 10299
Qy        829 CACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCC 888
Db      10300 CACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCC 10359
Qy        889 CCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 948
Db      10360 CCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 10419
Qy        949 CAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 1008
Db      10420 CAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 10479
Qy       1009 CTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCT 1068
Db      10480 CTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCT 10539
Qy       1069 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACC 1128
Db      10540 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCGCACC 10599
Qy       1129 CTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAG 1188
Db      10600 CTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCTCCAG 10659
Qy       1189 GAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAA 1248
Db      10660 GAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAA 10719
Qy       1249 GCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGG 1308
Db      10720 GCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGG 10779
```

```
Qy       1309 GTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCA 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10780 GTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCA 10839

Qy       1369 GCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10840 GCACACAAGCACATTGGGGCACCTGGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTT 10899

Qy       1429 CAGATCCTGGGGGAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10900 CAGATCCTGGGGGAGCCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGT 10959

Qy       1489 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATG 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10960 GCACCTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATG 11019

Qy       1549 GAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCT 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11020 GAGTCTCGTTCTGTCGCCCAGGCTGAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCT 11079

Qy       1609 CCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAG 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11080 CCACTCCCTGGGGCTCAAGCGATCCTCCCACCTCAGCCGCCGAAGTAGCTGGGACTATAG 11139

Qy       1669 GTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATG 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11140 GTGTGTACCATCACACCTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATG 11199

Qy       1729 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAA 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11200 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCCTCGGCCTCCCAAA 11259

Qy       1789 GTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCT 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11260 GTACTGGGATTACACGCAGAAGGCACCATGCCCAGGCTAGATGTGTCTTATCCCAATCCT 11319

Qy       1849 TTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11320 TTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 11379

Qy       1909 GGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAA 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11380 GGGTTAAGAGCCAGATAAGGAGAAATCCCTTTCCTAGGTTTGGAATGTGTTGTGAAAAAA 11439

Qy       1969 AAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGAC 2028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11440 AAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGAC 11499

Qy       2029 ATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAA 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11500 ATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAA 11559

Qy       2089 CTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTA 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11560 CTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTA 11619

Qy       2149 GGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCA 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11620 GGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCA 11679

Qy       2209 TGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACAT 2268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11680 TGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACAT 11739

Qy       2269 ATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCT 2328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11740 ATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCT 11799

Qy       2329 GCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATG 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11800 GCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATG 11859

Qy       2389 CGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAG 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11860 CGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAG 11919

Qy       2449 TCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCA 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11920 TCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCA 11979

Qy       2509 GGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11980 GGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 12039

Qy       2569 TCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTT 2628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12040 TCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATTAAACTGCAGTATTT 12099

Qy       2629 TGGAAAGCAAA 2639
             |||||||||| |
Db      12100 TGGAAAGCACA 12110
```

RESULT 4
US-09-764-868-493
; Sequence 493, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 493
; LENGTH: 1316
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1281)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1299)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-493

Query Match 48.1%; Score 1286; DB 9; Length 1316;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1286; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

```
Qy         45 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCC 60

Qy        105 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGC 120

Qy        165 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCC 180

Qy        225 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 284
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGTCGGCATAGACTTCAGSAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTG 240

Qy        285 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGA 300

Qy        345 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 GATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATC 360

Qy        405 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTA 420

Qy        465 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTG 480

Qy        525 GCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 584
              |||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db        481 GCCAGGGAGTACGGKGTTCCCTTYCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAG 540

Qy        585 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAG 600

Qy        645 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 704
              ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        601 CCCAGCTTCCAGATCCGAGACTATGTAKAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCC 660

Qy        705 TTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCC 764
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        661 TTCATGTGAATCCCAGGGGGCAKAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCC 720

Qy        765 CCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTC 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTC 780

Qy        825 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGT 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGT 840

Qy        885 CTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTA 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTACATAATTTA 900

Qy        945 ATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTG 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTG 960

Qy       1005 CTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCC 1064
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTTCTAGGACTTGGGSGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGTGTTGCTCC 1020

Qy       1065 AGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCG 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGACTCCATCG 1080

Qy       1125 CACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCT 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CACCCTCAGTTCAGCTGGACAGAGGCTCAGGTGACCCCAGCCTTCACTGTCTCCCGCTCT 1140

Qy       1185 CCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGAC 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCAGGAGCTTATCTTCGCCCCATCTCCCAAATAAGTGGGCCCTTGTGCTGTGAGGAAGAC 1200

Qy       1245 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 CAAAGCCTCAGGGAAGATAAGAGATATGGAGATGGGAGGGGGAGGACAAGGGGCAGAGAG 1260

Qy       1305 TAGGGTCTAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAG 1350
              ||||||:|||||||:|||||  ||| |||: | |   || ||  | |
Db        1261 TAGGGTYTAGCTGGYTATCTNTGGGCTTWATTAACACCNCCCTGGG 1306


RESULT 5
US-10-037-270-959
; Sequence 959, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 959
;   LENGTH: 1106
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)..(699)
US-10-037-270-959

  Query Match             41.1%;  Score 1099.6;  DB 9;  Length 1106;
  Best Local Similarity   99.6%;  Pred. No. 7e-295;
  Matches 1102;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 GGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 74
              || | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGTCCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGG 60

Qy         75 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAG 120

Qy        135 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGAC 180

Qy        195 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240

Qy        255 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGG 300

Qy        315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGAC 360

Qy        375 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTAT 420

Qy        435 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGA 480

Qy        495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 554
```

```
Db        481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAG 540

Qy        555 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 614

Db        541 ACCAGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTG 600

Qy        615 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 674

Db        601 AAATACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAG 660

Qy        675 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 734

Db        661 TCCCAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAG 720

Qy        735 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 794

Db        721 GCTCTGGAGGCACACAGGATGCAGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGC 780

Qy        795 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 854

Db        781 TGAGCCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 840

Qy        855 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 914

Db        841 TACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACAC 900

Qy        915 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 974

Db        901 TTAGCTTTTATTTTAATAGTACATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAA 960

Qy        975 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1034

Db        961 AACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCC 1020

Qy       1035 TCCTAAGCATAACAAAGGTGGTGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1094

Db       1021 TCCTAAGCATAACAAAGGTGGGGTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTT 1080

Qy       1095 TGGGGGTGAGGGCAGGTAATGACTCC 1120

Db       1081 TGGGGGTGAGGGCAGGTAATGACTCC 1106


RESULT 6
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 1116
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)...(699)
US-09-794-257-13

  Query Match            40.6%;  Score 1086;  DB 10;  Length 1116;
  Best Local Similarity  99.5%;  Pred. No. 4.2e-291;
  Matches 1089;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         37 GGGACATGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 96

Db         23 GCGTCCGGACGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCT 82

Qy         97 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 156

Db         83 CCCCGCCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAG 142

Qy        157 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 216

Db        143 GCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCT 202

Qy        217 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 276

Db        203 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAG 262

Qy        277 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 336

Db        263 TGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTT 322

Qy        337 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 396

Db        323 ATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCG 382

Qy        397 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 456

Db        383 ACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCA 442

Qy        457 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 516

Db        443 TGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAG 502

Qy        517 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 576

Db        503 AGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGA 562

Qy        577 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 636

Db        563 ATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGG 622

Qy        637 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 696

Db        623 CGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCT 682

Qy        697 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 756

Db        683 GCTGCTCCTTCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGC 742

Qy        757 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 816

Db        743 AGCCTTCCCCCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATG 802

Qy        817 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 876

Db        803 GAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTT 862

Qy        877 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 936

Db        863 CCTGCGGTCTCCCCGCATCCACAGGGAGGGTAAAACACTTAGCTTTTATTTTAATAGTAC 922

Qy        937 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 996

Db        923 ATAATTTAATACCAAAAAAGGCGCCTGGATCCCCAAAAAACCGAGGCTGGGAGCTAGTGG 982

Qy        997 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1056

Db        983 CCCTTTTGCTTTCTAGGACTTGGGGGGCCGGCCCTCCCTCCTAAGCATAACAAAGGTGGT 1042

Qy       1057 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAATGA 1116

Db       1043 GTTGCTCCAGCTCAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAAGTAATGA 1102

Qy       1117 CTCCATCGCACCCT 1130
```

```
Db       1103 CTCCATCGCACCCT 1116


RESULT 7
US-10-051-986-10
; Sequence 10, Application US/10051986
; Patent No. US20020146770A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Tang, Y. Tom
;                  Lal, Preeti
;                  Guegler, Karl J.
;                  Corley, Neil C.
;                  Patterson, Chandra
;                  Batra, Sajeev
;                  Baughn, Mariah R.
;       TITLE OF INVENTION: RAS PROTEINS
;       NUMBER OF SEQUENCES: 14
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;            STREET: 3174 Porter Drive
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: US
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/051,986
;            FILING DATE: 15-Jan-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/766,551
;            FILING DATE: DECEMBER 12, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cerrone, Michael C.
;            REGISTRATION NUMBER: 39,132
;            REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 10:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 875 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: UCMCL5T01
;            CLONE: 1528559
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10 :
US-10-051-986-10

  Query Match             32.6%;  Score 873;  DB 12;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-232;
  Matches  873;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCACGCCAGGCGCCGTTGCCACCCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCT 60

Qy        106 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCA 120

Qy        166 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAACATGTTTCCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCA 180

Qy        226 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC 240

Qy        286 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGATCTGGGACACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAG 300

Qy        346 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGCTCAGGCCTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCA 360

Qy        406 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTGGCTCACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAG 420

Qy        466 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAACAAGGCGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGG 480

Qy        526 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGGGAGTACGGTGTTCCCTTCCTGGAGACCAGCGCCAAGACTGGCATGAATGTGGAGT 540

Qy        586 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGCCTTTCTGGCCATCGCCAAGGAACTGAAATACCGGGCCGGGCATCAGGCGGATGAGC 600

Qy        646 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAGAAGAAGCGCTCCAGCTGCTGCTCCT 660

Qy        706 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCATGTGAATCCCAGGGGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCC 720

Qy        766 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTCCCAGGCCTGGCTTATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTCA 780

Qy        826 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGCACAGCCGCTTCCTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTC 840

Qy        886 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 918
              |||||||||||||||||||||||||||||||||
Db        841 TCCCCGCATCCACAGGGAGGGTAAAACACTTAG 873


RESULT 8
US-10-274-968-3
; Sequence 3, Application US/10274968
; Publication No. US20030049793A1
; GENERAL INFORMATION:
;  APPLICANT: YE, Jane et al
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR
;  TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER
;  TITLE OF INVENTION:  COFACTOR PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001052 DIV
;  CURRENT APPLICATION NUMBER: US/10/274,968
;  CURRENT FILING DATE:  2002-10-22
;  PRIOR APPLICATION NUMBER: 09/740,027
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 24707
;   TYPE: DNA
;   ORGANISM: Human
US-10-274-968-3

  Query Match             25.6%;  Score 683.4;  DB 9;  Length 24707;
  Best Local Similarity   99.9%;  Pred. No. 1.4e-178;
```

```
  Matches  684;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1955 GTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGC 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGC 60

Qy       2015 AAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCAT 120

Qy       2075 GCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCT 2134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCT 180

Qy       2135 ACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGA 2194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGA 240

Qy       2195 GTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC 2254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC 300

Qy       2255 CTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACA 2314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACA 360

Qy       2315 TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGG 2374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGG 420

Qy       2375 GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT 2434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT 480

Qy       2435 CTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTC 2494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTC 540

Qy       2495 CCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 2554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 600

Qy       2555 TGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATT 2614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATT 660

Qy       2615 AAACTGCAGTATTTTGGAAAGCAAA 2639
             ||||||||||||||||||||||| |
Db        661 AAACTGCAGTATTTTGGAAAGCACA 685
```

```
RESULT 9
US-09-740-027-3
; Sequence 3, Application US/09740027
; Patent No. US20020076749A1
; GENERAL INFORMATION:
;  APPLICANT: YE, Jane et al
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR
;  TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER
;  TITLE OF INVENTION:  COFACTOR PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001052
;  CURRENT APPLICATION NUMBER: US/09/740,027
;  CURRENT FILING DATE:  2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 24707
;   TYPE: DNA
;   ORGANISM: Human
US-09-740-027-3

  Query Match           25.6%;  Score 683.4;  DB 10;  Length 24707;
  Best Local Similarity  99.9%;  Pred. No. 1.4e-178;
```

```
  Matches  684;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1955 GTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGC 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGTTGTGAAAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGAGACAAGATTAAGC 60

Qy       2015 AAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAACCTCCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCCTCCCTGACCACCCAT 120

Qy       2075 GCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCT 2134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTACCTCCCCTCCT 180

Qy       2135 ACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGA 2194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGA 240

Qy       2195 GTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC 2254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAGTTCATGTGTC 300

Qy       2255 CTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACA 2314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACA 360

Qy       2315 TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGG 2374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGG 420

Qy       2375 GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT 2434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAAT 480

Qy       2435 CTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTC 2494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTGGCCTC 540

Qy       2495 CCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 2554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 600

Qy       2555 TGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATT 2614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTTTTTCAGAATT 660

Qy       2615 AAACTGCAGTATTTTGGAAAGCAAA 2639
             ||||||||||||||||||||||| |
Db        661 AAACTGCAGTATTTTGGAAAGCACA 685
```

```
RESULT 10
US-09-817-199A-24
; Sequence 24, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001187
;  CURRENT APPLICATION NUMBER: US/09/817,199A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-817-199A-24

  Query Match           22.5%;  Score 600.6;  DB 10;  Length 601;
  Best Local Similarity  99.8%;  Pred. No. 1.8e-156;
```

```
  Matches  600;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1983 GCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGA 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCCTGGAGCTGGTGGGAGACAAGATTAAGCAAACCTCCCCTGACATGTATCCCTTTGA 60

Qy       2043 CCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGAT 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCAAGCTCTGCCTCCTCCCTGACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGAT 120

Qy       2103 ACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTC 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCAGGGCCTAAACTGCTTTACCTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTC 180

Qy       2163 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTG 2222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTG 240

Qy       2223 GGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTC 2282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGTAGAGAGAGGGGTAGCAAGTTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTC 300

Qy       2283 TGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTAC 2342
             :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTAC 360

Qy       2343 TGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCT 2402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCAGCTCTGCCGGCCTGGAGGGGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCT 420

Qy       2403 GAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 2462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 480

Qy       2463 TTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGAC 2522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGGCCCGATAAGGAAGTTCTCCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGAC 540

Qy       2523 ATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCT 2582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTGCCAGTCTCTTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCT 600

Qy       2583 C 2583
              |
Db        601 C 601


RESULT 11
US-09-794-257-15
; Sequence 15, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel
;  TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
;  TITLE OF INVENTION:  Human G-Proteins
;  FILE REFERENCE: 35800/209285
;  CURRENT APPLICATION NUMBER: US/09/794,257
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER: 60/185,606
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 576
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-794-257-15

  Query Match            21.5%;  Score 576;  DB 10;  Length 576;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-149;
  Matches  576;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        138 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGG 60

Qy        198 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120

Qy        258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTC 180

Qy        318 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTATGACATC 240

Qy        378 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGAGTATGCC 300

Qy        438 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTG 360

Qy        498 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCCTTCCTGGAGACC 420

Qy        558 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCGCCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCATCGCCAAGGAACTGAAA 480

Qy        618 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCGGGCCGGGCATCAGGCGGATGAGCCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

Qy        678 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 713
             ||||||||||||||||||||||||||||||||||||
Db        541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATGTGA 576


RESULT 12
US-09-817-199A-25
; Sequence 25, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al.
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001187
;  CURRENT APPLICATION NUMBER: US/09/817,199A
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-817-199A-25

  Query Match            21.4%;  Score 573;  DB 10;  Length 601;
  Best Local Similarity  99.7%;  Pred. No. 8.3e-149;
  Matches  573;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2065 GACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTAC 2124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACCACCCATGCCCTTTCCTTTAACTTCTCAAACAGATACCAGGGCCTAAACTGCTTTAC 60

Qy       2125 CTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAA 2184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCCTCCTACTGAGTCAGGTTAGGTGGTGGGAGGTCACCCATTTCCGAGTTAAACCAA 120

Qy       2185 TGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAG 2244
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCAATATGAGTAAAACAAAGTCATGTGGGTATGTCTGGGGTAGAGAGAGGGGTAGCAAG 180

Qy       2245 TTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGG 2304
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCATGTGTCCTCCTTGGTCACATATCTCCCAAAGCTCTGATCCCTGCCATGGGAAGTGG 240

Qy       2305 ACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGG 2364
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACAGGAAACATGAGGTCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGGCCTGGAGG 300

Qy       2365 GGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCA 2424
                 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 RGGAGAGGGGGAGGAAGAAGTATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCA 360

Qy       2425 AGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCT 2484
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCGATAAGGAAGTTCT 420

Qy       2485 CCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCC 2544
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGTGGCCTCCCTCAGGCAGAGCAGGGAGGAGGCTGACATTGCCAGTCTCTTCTGGGGCC 480

Qy       2545 CAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTT 2604
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCTCTGGGCCTCTTGCATTTGAGTT 540

Qy       2605 TTTCAGAATTAAACTGCAGTATTTTGGAAAGCAAA 2639
                 |||||||||||||||||||||||||||||||| |
Db        541 TTTCAGAATTAAACTGCAGTATTTTGGAAAGCACA 575


RESULT 13
US-09-867-550-1811
; Sequence 1811, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1811
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)
;   OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1811

  Query Match             16.6%;  Score 443.4;  DB 10;  Length 447;
  Best Local Similarity   99.8%;  Pred. No. 6.8e-113;
  Matches  444;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 GGCCGGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCAC 70
                 |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGCCAGCACTGCTCACCTCTCGTCCAGGGACATGACGGGCACGCCAGGCGCCGTTGCCAC 62

Qy         71 CCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGG 130
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CCGGGATGGCGAGGCCCCCGAGCGCTCCCCGCCCTGCAGTCCGAGCTACGACCTCACGGG 122

Qy        131 CAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAA 190
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCTGATCCAATTCAA 182

Qy        191 AGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAA 250
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAA 242

Qy        251 GGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGA 310
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCAGGA 302

Qy        311 ACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTA 370
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCTTGCTTCTGCTGTA 362

Qy        371 TGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGA 430
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCTCACTGAGATTCATGA 422

Qy        431 GTATGCCCAGAGGGACGTGGTGATC 455
                 ||||||||||||||||||||||||
Db        423 GTATGCCCAGAGGGACGTGGTGATC 447


RESULT 14
US-09-918-995-34999
; Sequence 34999, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34999
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-34999

  Query Match             15.5%;  Score 415;  DB 9;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 5e-105;
  Matches  415;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1213 AAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATG 1272
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAATAAGTGGGCCCTTGTGCTGTGAGGAAGACCAAAGCCTCAGGGAAGATAAGAGATATG 62

Qy       1273 GAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTT 1332
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GAGATGGGAGGGGGAGGACAAGGGGCAGAGAGTAGGGTCTAGCTGGCTATCTCTGGCCTT 122

Qy       1333 ACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCT 1392
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ACTAACACCCCCCTGGAGGCATGCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCT 182

Qy       1393 GGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCC 1452
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGAAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGGAGCCCCTCCCC 242

Qy       1453 CCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAA 1512
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCCCTGAATCCCTGGCTTAGCTACCTTCCTGCCTGTGCACCTAAAAACCTCAGGTCAGAA 302

Qy       1513 CTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCT 1572
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CTAGGAAAAGAGTTTTGTTTTTATTTTTTTGAAATGGAGTCTCGTTCTGTCGCCCAGGCT 362
```

```
Qy       1573 GAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAG 1627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GAGGTGCAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAG 417


RESULT 15
US-09-918-995-648
; Sequence 648, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 648
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(434)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-648

  Query Match             8.7%;  Score 233.4;  DB 9;  Length 434;
  Best Local Similarity   77.7%;  Pred. No. 1.4e-54;
  Matches  282;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
              ||||||| || |   |||||| |||||  |||| ||| | || || || || || ||| |
Db         66 CTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGGACTCGGGTGTGGGGAAGACCTGTCT 125

Qy        176 CCTGATCCAATTCAAAGACGGGGCCTTCCTGTCCGGAACCTTCATAGCCACCGTCGGCAT 235
              ||| | | |||||| || || || |||||| | || ||||||||  ||||||| |||||
Db        126 GCTGGTGCGATTCAAGGATGGTGCTTTCCTGGCGGGGACCTTCATCTCCACCGTAGGCAT 185

Qy        236 AGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGA 295
              |||||| ||||||| ||  ||   |||||||| |||| ||||||||||||| |||||
Db        186 TGACTTCCGGAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGGA 245

Qy        296 CACCGCTGGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGC 355
              ||| ||||| ||||| |||||||| || || |||||||| || ||| | |||||||| ||
Db        246 CACAGCTGGTCAGGAGCGGTTCCGCAGTGTTACCCATGCCTACTACCGGGATGCTCATGC 305

Qy        356 CTTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAACATCAGGGCCTGGCT 415
              ||||  ||||| || ||  ||||||||||  | || || |||||||||  |||||||||
Db        306 TCTGCTGCTGCTCTACGATGTCACCAACAAGGCCTCCTTTGACAACATCCAGGCCTGGCT 365

Qy        416 CACTGAGATTCATGAGTATGCCCAGAGGGACGTGGTGATCATGCTGCTAGGCAACAAGGC 475
              || ||||| || ||||| ||||||   ||||||| | |||||||||| || |||||||
Db        366 GACCGAGATCCACGAGTACGCCCAGCACGACGTGGCGCTCATGCTGCTGGGGAACAAGGT 425

Qy        476 GGA 478
              |||
Db        426 GGA 428


Search completed: June 24, 2003, 02:35:58
Job time : 378 secs
```